Protein Sequence Searches - February 2005

All of the sequence databases on APSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension **.rup)** that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

  http://www.pir.uniprot.org/database/archive.shtml

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-1
Perfect score: 74
Sequence: 1 ATQKITYRISGVGID 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	37.8	10	2	A60589	sperm-activating p
2	27	36.5	10	2	E60787	sperm-activating p
3	25	33.8	15	2	A27504	histone H2A - mous
4	24	32.4	10	2	E60589	sperm-activating p
5	23	31.1	10	2	PC2044	beta-Kirilowin - M
6	22	29.7	10	2	PT0230	Ig heavy chain CDR
7	22	29.7	15	2	PA0075	fructose-bisphosph
8	22	29.7	15	2	PA0102	fructose-bisphosph
9	22	29.7	15	2	B45115	peptidylprolyl iso
10	21	28.4	10	2	PH0946	T-cell receptor be
11	21	28.4	10	2	G60787	sperm-activating p
12	21	28.4	10	2	F60589	sperm-activating p
13	21	28.4	10	2	G60589	sperm-activating p
14	21	28.4	14	2	PT0259	Ig heavy chain CRD
15	21	28.4	15	2	I65478	c-Ki-ras - hamster
16	21	28.4	15	2	A35417	28K serine protein
17	20	27.0	8	2	F60588	sperm-activating p
18	20	27.0	8	2	G60588	sperm-activating p
19	20	27.0	10	2	C45474	thrombospondin 2 -
20	20	27.0	10	2	A60588	sperm-activating p
21	20	27.0	10	2	E39572	sperm-activating p
22	20	27.0	10	2	D60589	sperm-activating p
23	20	27.0	10	2	B60589	sperm-activating p
24	20	27.0	12	2	A33099	163K exoantigen -
25	20	27.0	12	2	S29830	dimethylaniline mo
26	20	27.0	14	2	C35141	T-cell receptor de
27	20	27.0	15	2	PA0097	starch phosphoryla
28	20	27.0	15	2	S70719	H+-transporting tw
29	19.5	26.4	12	2	S21163	NAD ADP-ribosyltra
30	19	25.7	8	2	S63493	dissimilatory sulf
31	19	25.7	8	2	S65647	2-hydroxyglutaryl-
32	19	25.7	9	2	I49406	bone gla protein -
33	19	25.7	10	2	A47364	placental lactogen
34	19	25.7	10	2	F60787	sperm-activating p
35	19	25.7	10	2	I60588	sperm-activating p
36	19	25.7	12	2	D28551	hypothetical prote
37	19	25.7	13	2	S32473	lymnaDFamide 3 - g
38	19	25.7	14	2	S59495	formate dehydrogen
39	19	25.7	15	2	PH1320	Ig heavy chain DJ
40	19	25.7	15	2	S71396	dihydropyrimidine
41	18.5	25.0	15	2	PT0095	H+-transporting tw
42	18	24.3	6	2	A41946	T-cell receptor ga
43	18	24.3	10	2	PT0322	Ig heavy chain CRD
44	18	24.3	10	2	C60527	sperm-activating p
45	18	24.3	10	2	G60527	sperm-activating p

ALIGNMENTS

RESULT 1
A60589
sperm-activating peptide (Tyr-2, Asn-3, Gly-5, Ile-9, Asp-10 SAP-I) - slate-pencil urchi
C;Species: Heterocentrotus mamillatus
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C;Accession: A60589
R;Yoshino, K.I.; Kajiura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamagu
Comp. Biochem. Physiol. B 94, 739-751, 1989
A;Title: A halogenated amino acid-containing sperm activating peptide and its related p
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A;Reference number: A60527
A;Accession: A60589
A;Molecule type: protein
A;Residues: 1-10 <YOS>
A;Cross-references: UNIPROT:Q7M4B7
C;Superfamily: unassigned animal peptides

Query Match 37.8%; Score 28; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 76;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 YRISGVGID 15
              | : | |||
Db          2 YNLGGGGID 10
```

RESULT 2
E60787
sperm-activating peptide (Ser-3,5,7, Asp-10 speract) - sea urchin (Hemicentrotus pulche
C;Species: Hemicentrotus pulcherrimus
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 16-Aug-2004
C;Accession: E60787
R;Suzuki, N.; Kajiura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka,
Comp. Biochem. Physiol. B 89, 687-693, 1988
A;Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentr
A;Reference number: A60787; MUID:88242184; PMID:3378407
A;Accession: E60787
A;Molecule type: protein
A;Residues: 1-10 <SUZ>
A;Cross-references: UNIPROT:Q7M4C9
C;Comment: This oligopeptide from egg jelly is one of several from this species, all of
at shows some, but not absolute, species restriction.

Query Match 36.5%; Score 27; DB 2; Length 10;
Best Local Similarity 44.4%; Pred. No. 1.2e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy          7 YRISGVGID 15
              : :|| |:|
Db          2 PSLSGSGVD 10
```

RESULT 3
A27504
histone H2A - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 17-Feb-1994
C;Accession: A27504
R;Liu, T.J.; Liu, L.; Marzluff, W.F.
Nucleic Acids Res. 15, 3023-3039, 1987
A;Title: Mouse histone H2A and H2B genes: four functional genes and a pseudogene undergo
A;Reference number: A93660; MUID:87174824; PMID:3562244
A;Accession: A27504
A;Molecule type: DNA
A;Residues: 1-15 <LIU>

Query Match 33.8%; Score 25; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 TYRISGVG 13
|| :|| |
Db 4 TYMMSGRG 11

RESULT 4
E60589
sperm-activating peptide (Tyr-2, Ser-3,5, Ala-8, Asp-10 SAP-I) - Echinometra mathaei
C;Species: Echinometra mathaei
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Aug-2004
C;Accession: E60589
R;Yoshino, K.I.; Kajiura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamaguc
Comp. Biochem. Physiol. B 94, 739-751, 1989
A;Title: A halogenated amino acid-containing sperm activating peptide and its related pe
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A;Reference number: A60527
A;Accession: E60589
A;Molecule type: protein
A;Residues: 1-10 <YOS>
A;Cross-references: UNIPROT:Q7M4C1
A;Note: an identical peptide was isolated from Echinometra mathaei type A and type B

Query Match 32.4%; Score 24; DB 2; Length 10;
Best Local Similarity 44.4%; Pred. No. 4e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 7 YRISGVGID 15
| :|| :|
Db 2 YSLSGGAVD 10

RESULT 5
PC2044
beta-Kirilowin - Mongolian snake-gourd (fragment)
C;Species: Trichosanthes kirilowii (Mongolian snake-gourd)
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: PC2044
R;Dong, T.X.; Ng, T.B.; Yeung, H.W.; Wong, R.N.S.
Biochem. Biophys. Res. Commun. 199, 387-393, 1994
A;Title: Isolation and characterization of a novel ribosome-inactivating protein, beta-k
A;Reference number: PC2044; MUID:94168605; PMID:8123040
A;Accession: PC2044
A;Molecule type: protein
A;Residues: 1-10 <DON>
A;Cross-references: UNIPROT:Q7M1I6
A;Experimental source: seed
C;Comment: This protein exhibited strong abortifacient activity, and is a ribosome inact
C;Keywords: seed

Query Match 31.1%; Score 23; DB 2; Length 10;
Best Local Similarity 57.1%; Pred. No. 6.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 ITYRISG 11
| :|:||
Db 3 IIFRLSG 9

RESULT 6
PT0230
Ig heavy chain CDR3 region (clone 1-118A) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0230
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0230
A;Molecule type: DNA
A;Residues: 1-10 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 29.7%; Score 22; DB 2; Length 10;
Best Local Similarity 83.3%; Pred. No. 9.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 ISGVGI 14
| ||||
Db 5 IFGVGI 10

RESULT 7
PA0075
fructose-bisphosphate aldolase (EC 4.1.2.13) I - fungus (Fusarium sporotrichioides) (fra
N;Alternate names: aldolase; fructose-1,6,-biophosphate triosephosphate-lase
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0075; PA0077
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0075
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Z5
A;Note: this form (I) had a molecular weight of 30.6K and an isoelectric point of 5.3
A;Accession: PA0077
A;Molecule type: protein
A;Residues: 1-15 <CH2>
A;Note: this form (II) had a molecular weight of 31.6K and an isoelectric point of 5.4
C;Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 29.7%; Score 22; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.4e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 QKITYRISGV 12
|:: | |||
Db 3 QEVLSRDSGV 12

RESULT 8
PA0102
fructose-bisphosphate aldolase (EC 4.1.2.13) III - fungus (Fusarium sporotrichioides) (f
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0102
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0102
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Z4

```
C;Keywords: aldehyde-lyase; carbon-carbon lyase

  Query Match             29.7%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 QKITYRISGV 12
              |::  | |||
Db          3 QEVLSRKSGV 12


RESULT 9
B45115
peptidylprolyl isomerase (EC 5.2.1.8) FKBP51 - human (fragment)
N;Alternate names: FK506-binding protein FKBP51; peptidylprolyl cis-trans isomerase FKBP
C;Species: Homo sapiens (man)
C;Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B45115
R;Wiederrecht, G.; Hung, S.; Chan, H.K.; Marcy, A.; Martin, M.; Calaycay, J.; Boulton, D
J. Biol. Chem. 267, 21753-21760, 1992
A;Title: Characterization of high molecular weight FK-506 binding activities reveals a n
A;Reference number: A45115; MUID:93016131; PMID:1383226
A;Accession: B45115
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <WIE>
A;Cross-references: UNIPROT:Q9UDK1
A;Experimental source: JURKAT cells
A;Note: sequence extracted from NCBI backbone (NCBIP:116748)
C;Keywords: cis-trans-isomerase; cyclosporin A binding

  Query Match             29.7%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 ATQKITY 7
              ||: | |
Db          1 ATESIAY 7


RESULT 10
PH0946
T-cell receptor beta chain V-D-J region (clone BB12) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C;Accession: PH0946
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergi
A;Reference number: PH0891; MUID:92078857; PMID:1836012
A;Accession: PH0946
A;Molecule type: mRNA
A;Residues: 1-10 <GOL>
A;Experimental source: myelin basic protein fragment-reactive T-cell, recovered from exp
C;Keywords: T-cell receptor

  Query Match             28.4%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         10 SGVGID 15
              || | |
Db          4 SGTGTD 9


RESULT 11
G60787
sperm-activating peptide (Ser-3,5,7 speract) - sea urchin (Anthocidaris crassispina)
C;Species: Anthocidaris crassispina
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 16-Aug-2004
C;Accession: G60787
R;Suzuki, N.; Kajiura, H.; Nomura, K.; Garbers, D.L.; Yoshino, K.; Kurita, M.; Tanaka, H
Comp. Biochem. Physiol. B 89, 687-693, 1988
A;Title: Some more speract derivatives associated with eggs of sea urchins, Pseudocentro
A;Reference number: A60787; MUID:88242184; PMID:3378407
A;Accession: G60787
A;Molecule type: protein
A;Residues: 1-10 <SUZ>
A;Cross-references: UNIPROT:Q7M4D8
C;Comment: This oligopeptide from egg jelly is one of several from this species, all of
at shows some, but not absolute, species restriction.

  Query Match             28.4%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   37.5%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 YRISGVGI 14
              : :|| |:
Db          2 FSLSGSGV 9


RESULT 12
F60589
sperm-activating peptide (Asn-3, Ser-5 SAP-I) - Echinometra mathaei (type A)
C;Species: Echinometra mathaei
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Aug-2004
C;Accession: F60589
R;Yoshino, K.I.; Kajiura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamaguc
Comp. Biochem. Physiol. B 94, 739-751, 1989
A;Title: A halogenated amino acid-containing sperm activating peptide and its related pe
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A;Reference number: A60527
A;Accession: F60589
A;Molecule type: protein
A;Residues: 1-10 <YOS>
A;Cross-references: UNIPROT:Q7M4C5

  Query Match             28.4%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   37.5%;  Pred. No. 1.4e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 YRISGVGI 14
              : :|| |:
Db          2 FNLSGGGV 9


RESULT 13
G60589
sperm-activating peptide (Tyr-2, Asn-3, Asp-7,10, Arg-8, Ile-9 SAP-I) - Echinometra math
C;Species: Echinometra mathaei
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 16-Aug-2004
C;Accession: G60589
R;Yoshino, K.I.; Kajiura, H.; Nomura, K.; Takao, T.; Shimonishi, Y.; Kurita, M.; Yamagu
Comp. Biochem. Physiol. B 94, 739-751, 1989
A;Title: A halogenated amino acid-containing sperm activating peptide and its related pe
otus nudus, Echinometra mathaei and Heterocentrotus mammillatus.
A;Reference number: A60527
A;Accession: G60589
A;Molecule type: protein
A;Residues: 1-10 <YOS>
A;Cross-references: UNIPROT:Q7M4C2

  Query Match             28.4%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   44.4%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 YRISGVGID 15
              | ::|  ||
Db          2 YNLNGDRID 10


RESULT 14
PT0259
Ig heavy chain CRD3 region (clone 2-118C) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
```

C;Accession: PT0259
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0259
A;Molecule type: DNA
A;Residues: 1-14 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 28.4%; Score 21; DB 2; Length 14;
Best Local Similarity 57.1%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 ISGVGID 15
|: |||
Db 6 IAAAGID 12

RESULT 15
I65478
c-Ki-ras - hamster (fragment)
C;Species: Cricetinae gen. sp. (hamster)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 28-Feb-1997
C;Accession: I65478
R;Takahashi, T.; Moyer, M.P.; Cano, M.; Wang, Q.J.; Mountjoy, C.P.; Sanger, W.; Adrian,
Carcinogenesis 16, 931-939, 1995
A;Title: Differences in molecular biological, biological and growth characteristics betw
A;Reference number: I52734; MUID:95246257; PMID:7728976
A;Accession: I65478
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-15 <RES>
A;Cross-references: GB:S77069; NID:g914177
C;Genetics:
A;Gene: c-Ki-ras

Query Match 28.4%; Score 21; DB 2; Length 15;
Best Local Similarity 44.4%; Pred. No. 2.2e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 7 YRISGVGID 15
|:: || |
Db 4 YKLVVVGAD 12

Search completed: February 22, 2005, 09:46:22
Job time : 12.1333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-1
Perfect score: 74
Sequence: 1 ATQKITYRISGVGID 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	37.8	10	2	Q7M4B7	Q7m4b7 heterocentr
2	27	36.5	10	2	Q7M4C9	Q7m4c9 hemicentrot
3	27	36.5	15	2	Q9QVN0	Q9qvn0 rattus sp.
4	26	35.1	12	1	OPS3_DROVI	P17645 drosophila
5	24	32.4	10	2	Q7M4C1	Q7m4c1 echinometra
6	24	32.4	11	2	O77899	O77899 oreochromis
7	24	32.4	11	2	O77900	O77900 oreochromis
8	24	32.4	11	2	O77901	O77901 oreochromis
9	24	32.4	11	2	O77902	O77902 oreochromis
10	24	32.4	11	2	O77903	O77903 oreochromis
11	24	32.4	11	2	O77904	O77904 oreochromis
12	24	32.4	11	2	O77905	O77905 oreochromis
13	24	32.4	11	2	O77916	O77916 oreochromis
14	24	32.4	11	2	O77917	O77917 oreochromis
15	24	32.4	11	2	O77921	O77921 pseudotroph
16	24	32.4	13	1	FLET_LIMFT	P82064 limnodynast
17	23	31.1	10	1	VEG6_BACSU	P80699 bacillus su
18	23	31.1	10	2	Q7M1I6	Q7m1i6 trichosanth
19	23	31.1	15	1	DIDH_PSESP	P80701 pseudomonas
20	23	31.1	15	2	Q9UCC0	Q9ucc0 homo sapien
21	22	29.7	15	2	Q7M4Z4	Q7m4z4 fusarium sp
22	22	29.7	15	2	Q7M4Z5	Q7m4z5 fusarium sp
23	21	28.4	8	1	DYS1_LIMIN	P82079 limnodynast
24	21	28.4	8	2	O52062	O52062 bacillus me
25	21	28.4	10	2	Q7M4C2	Q7m4c2 echinometra
26	21	28.4	10	2	Q7M4C5	Q7m4c5 echinometra
27	21	28.4	10	2	Q7M4D8	Q7m4d8 anthocidari
28	21	28.4	15	1	R13A_SPIOL	P82454 spinacia ol
29	21	28.4	15	2	Q6JDJ7	Q6jdj7 canis famil
30	21	28.4	15	2	Q7M3G3	Q7m3g3 bos taurus
31	21	28.4	15	2	Q9WTA3	Q9wta3 amorphophal
32	21	28.4	15	2	Q80X04	Q80x04 mesocricetu
33	20	27.0	8	2	Q7M4C6	Q7m4c6 pseudobolet
34	20	27.0	10	2	Q7M3T4	Q7m3t4 tripneustes
35	20	27.0	10	2	Q7M4B3	Q7m4b3 strongyloce
36	20	27.0	10	2	Q7M4B9	Q7m4b9 heterocentr
37	20	27.0	10	2	Q7M4C4	Q7m4c4 echinometra
38	20	27.0	10	2	Q95H99	Q95h99 papio anubi
39	20	27.0	10	2	Q95HF4	Q95hf4 papio anubi
40	20	27.0	10	2	Q95HF5	Q95hf5 papio anubi
41	20	27.0	10	2	Q95HF6	Q95hf6 papio anubi
42	20	27.0	10	2	Q95HF7	Q95hf7 papio anubi
43	20	27.0	10	2	Q95HF8	Q95hf8 papio anubi
44	20	27.0	10	2	Q95HF9	Q95hf9 papio anubi
45	20	27.0	10	2	Q95HG0	Q95hg0 papio anubi

ALIGNMENTS

RESULT 1
Q7M4B7
ID Q7M4B7 PRELIMINARY; PRT; 10 AA.
AC Q7M4B7;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sperm-activating peptide (Tyr-2, Asn-3, Gly-5, Ile-9, Asp-10 SAP-
DE I).
OS Heterocentrotus mammillatus (Slate-pencil urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Echinometridae;
OC Heterocentrotus.
OX NCBI_TaxID=31180;
RN [1]
RP SEQUENCE.
RA Yoshino K.I., Kajiura H., Nomura K., Takao T., Shimonishi Y.,
RA Kurita M., Yamaguchi M., Suzuki N.;
RT "A halogenated amino acid-containing sperm activating peptide and its
RT related peptides isolated from the egg jelly of sea urchins,
RT Tripneustes gratilla, Pseudoboletia maculata, Strongylocentrotus
RT nudus, Echinometra mathaei and Heterocentrotus mammillatus.";
RL Comp. Biochem. Physiol. 94:739-751(1989).
DR PIR; A60589; A60589.
SQ SEQUENCE 10 AA; 922 MW; 128AC37878787724 CRC64;

Query Match 37.8%; Score 28; DB 2; Length 10;
Best Local Similarity 55.6%; Pred. No. 7.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 YRISGVGID 15
              | : | |||
Db          2 YNLGGGGID 10
```

RESULT 2
Q7M4C9
ID Q7M4C9 PRELIMINARY; PRT; 10 AA.
AC Q7M4C9;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Sperm-activating peptide (Ser-3,5,7, Asp-10 speract).
OS Hemicentrotus pulcherrimus (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Hemicentrotus.
OX NCBI_TaxID=7650;
RN [1]
RP SEQUENCE.
RA Suzuki N., Kajiura H., Nomura K., Garbers D.L., Yoshino K., Kurita M.,
RA Tanaka H., Yamaguchi M.;
RT "Some more speract derivatives associated with eggs of sea urchins,
RT Pseudocentrotus depressus, Strongylocentrotus purpuratus,

```
RT   Hemicentrotus pulcherrimus and Anthocidaris crassispina.";
RL   Comp. Biochem. Physiol. 89:687-693(1988).
DR   PIR; E60787; E60787.
SQ   SEQUENCE   10 AA;  925 MW;  933B5365B865B735 CRC64;

  Query Match             36.5%;  Score 27;  DB 2;  Length 10;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 YRISGVGID 15
              : :|| |:|
Db          2 FSLSGSGVD 10


RESULT 3
Q9QVN0
ID   Q9QVN0      PRELIMINARY;      PRT;    15 AA.
AC   Q9QVN0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Glia-derived nexin, GDN, protease nexin-1 (Fragment).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92207980; PubMed=1554734;
RA   Rovelli G., Stone S.R., Guidolin A., Sommer J., Monard D.;
RT   "Characterization of the heparin-binding site of glia-derived
RT   nexin/protease nexin-1.";
RL   Biochemistry 31:3542-3549(1992).
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1702 MW;  3D999487E58AF9A2 CRC64;

  Query Match             36.5%;  Score 27;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 YRISGVG 13
              | ::|||
Db          2 YNVNGVG 8


RESULT 4
OPS3_DROVI
ID   OPS3_DROVI     STANDARD;      PRT;    12 AA.
AC   P17645;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Opsin Rh3 (Inner R7 photoreceptor cells opsin) (Fragment).
GN   Name=Rh3;
OS   Drosophila virilis (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7244;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90249748; PubMed=2140105;
RA   Fortini M.E., Rubin G.M.;
RT   "Analysis of cis-acting requirements of the Rh3 and Rh4 genes reveals
RT   a bipartite organization to rhodopsin promoters in Drosophila
RT   melanogaster.";
RL   Genes Dev. 4:444-463(1990).
CC   -!- FUNCTION: Visual pigments are the light-absorbing molecules that
CC       mediate vision. They consist of an apoprotein, opsin, covalently
CC       linked to cis-retinal.
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -!- MISCELLANEOUS: Each Drosophila eye is composed of 800 facets or
CC       ommatidia. Each ommatidium contains 8 photoreceptor cells (R1-R8),
CC       the R1 to R6 cells are outer cells, while R7 and R8 are inner
CC       cells.
CC   -!- MISCELLANEOUS: Opsin Rh3 is sensitive to UV light.
CC   -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC       Opsin subfamily.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X51350; CAA35742.1; -.
DR   FlyBase; FBgn0013091; Dvir\Rh3.
DR   InterPro; IPR000276; GPCR_Rhodpsn.
DR   InterPro; IPR001760; Opsin.
DR   PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; PARTIAL.
DR   PROSITE; PS00238; OPSIN; PARTIAL.
KW   G-protein coupled receptor; Glycoprotein; Phosphorylation;
KW   Photoreceptor; Retinal protein; Transmembrane; Vision.
FT   CARBOHYD     10     10       N-linked (GlcNAc. . .) (Probable).
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1253 MW;  04024E43495865B0 CRC64;

  Query Match             35.1%;  Score 26;  DB 1;  Length 12;
  Best Local Similarity   57.1%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 YRISGVG 13
              : |||:|
Db          3 FNISGIG 9


RESULT 5
Q7M4C1
ID   Q7M4C1      PRELIMINARY;      PRT;    10 AA.
AC   Q7M4C1;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Sperm-activating peptide (Tyr-2, Ser-3,5, Ala-8, Asp-10 SAP-I).
OS   Echinometra mathaei (Rock boring urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Echinometridae;
OC   Echinometra.
OX   NCBI_TaxID=31178;
RN   [1]
RP   SEQUENCE.
RA   Yoshino K.I., Kajiura H., Nomura K., Takao T., Shimonishi Y.,
RA   Kurita M., Yamaguchi M., Suzuki N.;
RT   "A halogenated amino acid-containing sperm activating peptide and its
RT   related peptides isolated from the egg jelly of sea urchins,
RT   Tripneustes gratilla, Pseudoboletia maculata, Strongylocentrotus
RT   nudus, Echinometra mathaei and Heterocentrotus mammillatus.";
RL   Comp. Biochem. Physiol. 94:739-751(1989).
DR   PIR; E60589; E60589.
SQ   SEQUENCE   10 AA;  925 MW;  8B9856D87865B735 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 10;
  Best Local Similarity   44.4%;  Pred. No. 4e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 YRISGVGID 15
              | :||  :|
Db          2 YSLSGGAVD 10


RESULT 6
O77899
ID   O77899      PRELIMINARY;      PRT;    11 AA.
```

```
AC   O77899;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050009; AAC41348.1; -.
FT   NON_TER      1      1
FT   NON_TER     11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 7
O77900
ID   O77900      PRELIMINARY;      PRT;    11 AA.
AC   O77900;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050010; AAC41349.1; -.
FT   NON_TER      1      1
FT   NON_TER     11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 8
O77901
ID   O77901      PRELIMINARY;      PRT;    11 AA.
AC   O77901;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050011; AAC41350.1; -.
FT   NON_TER      1      1
FT   NON_TER     11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 9
O77902
ID   O77902      PRELIMINARY;      PRT;    11 AA.
AC   O77902;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050012; AAC41351.1; -.
FT   NON_TER      1      1
FT   NON_TER     11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 10
O77903
ID   O77903      PRELIMINARY;      PRT;    11 AA.
AC   O77903;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
```

```
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050013; AAC41352.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 11
O77904
ID   O77904      PRELIMINARY;      PRT;    11 AA.
AC   O77904;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050014; AAC41353.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 12
O77905
ID   O77905      PRELIMINARY;      PRT;    11 AA.
AC   O77905;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050015; AAC41354.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 13
O77916
ID   O77916      PRELIMINARY;      PRT;    11 AA.
AC   O77916;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050029; AAC41368.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 14
O77917
ID   O77917      PRELIMINARY;      PRT;    11 AA.
AC   O77917;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
DE   MHC class II B locus 14 (Fragment).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050030; AAC41369.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


RESULT 15
O77921
ID   O77921      PRELIMINARY;      PRT;    11 AA.
AC   O77921;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   MHC class II B locus 14 (Fragment).
OS   Pseudotropheus sp. 'Pseudotropheus tropheops complex'.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Pseudotropheus.
OX   NCBI_TaxID=51796;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA   Figueroa F., Sultmann H., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid MHC
RT   class II B loci.";
RL   Genetics 149:1527-1537(1998).
DR   EMBL; AF050034; AAC41373.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1349 MW;  81C12D8EB7341B41 CRC64;

  Query Match             32.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 ITYRIS 10
              :|||:|
Db          3 MTYRLS 8


Search completed: February 22, 2005, 09:37:45
Job time : 53.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-1
Perfect score: 74
Sequence: 1 ATQKITYRISGVGID 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	74	100.0	15	2	AAW04841	Aaw04841 Self epit
2	30	40.5	14	4	AAM96744	Aam96744 Human pep
3	29	39.2	14	4	AAM98436	Aam98436 Human pep
4	29	39.2	14	4	ABB56775	Abb56775 Human SNP
5	28	37.8	6	2	AAR90448	Aar90448 Antibody
6	28	37.8	6	2	AAY55175	Aay55175 Anti CD34
7	28	37.8	6	2	AAY55183	Aay55183 Anti CD34
8	28	37.8	6	3	AAY86929	Aay86929 Human hae
9	28	37.8	6	3	AAY86921	Aay86921 Human hae
10	28	37.8	9	8	ADH40843	Adh40843 Human CD6
11	28	37.8	15	8	ADN65292	Adn65292 HLA bindi
12	27	36.5	10	4	AAG86324	Aag86324 Saccharom
13	27	36.5	10	4	AAG86188	Aag86188 Saccharom
14	27	36.5	10	8	ADN70236	Adn70236 Human 273
15	27	36.5	10	8	ADN67283	Adn67283 Human 273
16	27	36.5	10	8	ADN70105	Adn70105 Human 273
17	27	36.5	10	8	ADN67511	Adn67511 Human 273
18	27	36.5	11	8	ADQ91410	Adq91410 CDR3 of t
19	27	36.5	13	8	ADM12154	Adm12154 MHC class
20	27	36.5	13	8	ADO38372	Ado38372 Ara h 3 M
21	27	36.5	15	8	ADM12156	Adm12156 MHC class
22	27	36.5	15	8	ADN71740	Adn71740 Human 273
23	27	36.5	15	8	ADN70559	Adn70559 Human 273
24	27	36.5	15	8	ADN71499	Adn71499 Human 273
25	27	36.5	15	8	ADN71122	Adn71122 Human 273
26	27	36.5	15	8	ADN71721	Adn71721 Human 273
27	27	36.5	15	8	ADN70750	Adn70750 Human 273
28	27	36.5	15	8	ADN71576	Adn71576 Human 273
29	27	36.5	15	8	ADN70618	Adn70618 Human 273
30	27	36.5	15	8	ADO38374	Ado38374 Ara h 3 M
31	26.5	35.8	10	4	ABP22244	Abp22244 HIV A03 m
32	26.5	35.8	10	4	ABP14250	Abp14250 HIV A02 s
33	26.5	35.8	10	4	ABP16613	Abp16613 HIV A24 s
34	26.5	35.8	11	4	ABP14257	Abp14257 HIV A02 s
35	26.5	35.8	15	4	ABP25033	Abp25033 HIV DR 3b
36	26	35.1	7	8	ADP47064	Adp47064 Murine li
37	26	35.1	8	8	ADK09053	Adk09053 Human pap
38	26	35.1	9	6	ABP75214	Abp75214 Chlamydia
39	26	35.1	9	6	ABR63363	Abr63363 Human Alz
40	26	35.1	9	7	ABU64342	Abu64342 Human AAC
41	26	35.1	9	8	ADK09054	Adk09054 Human pap
42	26	35.1	10	2	AAR89616	Aar89616 Prostate
43	26	35.1	10	2	AAW33453	Aaw33453 Oligopept
44	26	35.1	10	2	ADG45903	Adg45903 Benign pr
45	26	35.1	10	5	AAE24261	Aae24261 Murine E-

ALIGNMENTS

RESULT 1
AAW04841
ID AAW04841 standard; peptide; 15 AA.
XX
AC AAW04841;
XX
DT 18-FEB-1997 (first entry)
XX
DE Self epitope of desmoglein 3, implicated in autoimmune disease.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Homo sapiens.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 1; Page 38; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is derived from the human desmoglein

```
CC   3 protein (amino acids 78-93 (sic)) and is implicated as a self epitope
CC   in pemphigus vulgaris. Peptides derived from the human desmoglein protein
CC   are described in AAW04841-47
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 74;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATQKITYRISGVGID 15
              |||||||||||||||
Db          1 ATQKITYRISGVGID 15


RESULT 2
AAM96744
ID   AAM96744 standard; peptide; 14 AA.
XX
AC   AAM96744;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human peptide #19 encoded by a SNP oligonucleotide.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US035498.
XX
PR   28-DEC-1999;   99US-0173419P.
PR   27-DEC-2000; 2000US-00173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT   autoimmune diseases and infections.
XX
PS   Disclosure; Page 3672; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC   encoding polymorphic variants of proteins related to amylases, amyloid
CC   proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,
CC   polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC   complement related proteins, cytochromes, kinesins, cytokines,
CC   interferons, interleukins, G-protein coupled receptors and thioesterases.
CC   The present sequence is a peptide encoded by one such oligonucleotide.
CC   The oligonucleotides and the peptides encoded by them may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate expression of the proteins listed above. Disorders that may
CC   be prevented, diagnosed and/or treated include multifactorial diseases
CC   with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC   arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus
CC   and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC   brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC   system and an infection of pathogenic organisms
XX
SQ   Sequence 14 AA;

  Query Match             40.5%;  Score 30;  DB 4;  Length 14;
  Best Local Similarity   55.6%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 TYRISGVGI 14
              |: | |:||
Db          1 TWEIQGIGI 9


RESULT 3
AAM98436
ID   AAM98436 standard; peptide; 14 AA.
XX
AC   AAM98436;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human peptide #1711 encoded by a SNP oligonucleotide.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US035498.
XX
PR   28-DEC-1999;   99US-0173419P.
PR   27-DEC-2000; 2000US-00173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT   autoimmune diseases and infections.
XX
PS   Disclosure; Page 4043; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC   encoding polymorphic variants of proteins related to amylases, amyloid
CC   proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,
CC   polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC   complement related proteins, cytochromes, kinesins, cytokines,
CC   interferons, interleukins, G-protein coupled receptors and thioesterases.
CC   The present sequence is a peptide encoded by one such oligonucleotide.
CC   The oligonucleotides and the peptides encoded by them may be used in the
CC   prevention, diagnosis and treatment of diseases associated with
CC   inappropriate expression of the proteins listed above. Disorders that may
CC   be prevented, diagnosed and/or treated include multifactorial diseases
CC   with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC   arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus
CC   and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC   brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC   system and an infection of pathogenic organisms
XX
SQ   Sequence 14 AA;
```

Query Match 39.2%; Score 29; DB 4; Length 14;
Best Local Similarity 62.5%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 KITYRISG 11
|:| |:||
Db 1 KLTVRVSG 8

RESULT 4
ABB56775
ID ABB56775 standard; peptide; 14 AA.
XX
AC ABB56775;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human SNP related amino acid sequence SEQ ID NO:1340.
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein.
XX
OS Homo sapiens.
XX
PN WO200138586-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US032311.
XX
PR 24-NOV-1999; 99US-0167383P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-355949/37.
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a pathology,
PT e.g. autoimmune diseases, ascribed to the presence of a sequence
PT polymorphism.
XX
PS Claim 1; Page 643; 674pp; English.
XX
CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the polymorphic
CC protein within appropriate physiological samples)
XX
SQ Sequence 14 AA;

Query Match 39.2%; Score 29; DB 4; Length 14;
Best Local Similarity 54.5%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 QKITYRISGVG 13
|: | |:|| |
Db 3 QEATARLSGSG 13

RESULT 5
AAR90448
ID AAR90448 standard; peptide; 6 AA.
XX
AC AAR90448;
XX
DT 03-SEP-1996 (first entry)
XX
DE Antibody 561 displacing peptide #16.
XX
KW Antibody; displacement; cell separation system; cell surface antigen; Ag;
KW random peptide display library; complementarity determining region; CDR;
KW antibody; Ab; peptide; CD34 cell; haematopoietic cell; tumour cell;
KW lymphocyte; high dose therapy; immune system; chemotherapy;
KW patient-specific vaccine.
XX
OS Synthetic.
XX
PN WO9534817-A1.
XX
PD 21-DEC-1995.
XX
PF 13-JUN-1995; 95WO-US007491.
XX
PR 14-JUN-1994; 94US-00259427.
XX
PA (BAXT) BAXTER INT INC.
XX
PI Tseng-Law J, Kobori JA, Al-Abdaly FA, Guillermo R, Helgerson SL;
PI Deans RJ;
XX
DR WPI; 1996-049806/05.
XX
PT Selecting target cells by reacting specific antibody to surface antigen -
PT then disrupting the complex formed with peptide displacer, partic. for
PT selecting CD34 cells for reconstitution of immune system after anticancer
PT treatment.
XX
PS Claim 25; Page 155; 170pp; English.
XX
CC AAR90413-R90468 represent antibody displacing peptides. These sequences
CC displace the monoclonal antibody designated 561. These sequences can be
CC used in a method to select target cells from a heterogeneous cell
CC suspension. In the suspension, there is at least one complex of a cell
CC separation system, linked to a primary antibody (Ab) which is bound to
CC the cell surface antigen (Ag). The complex is separated from the rest of
CC the suspension, and contacted with a peptide (such as this sequence) that
CC binds to the Ab, displacing it from the Ag and releasing the cell. These
CC sequences can be identified by biopanning random peptide display
CC libraries with the Ab, by analysis of potential antigenic peaks of the
CC cell surface antigen, or by analysis of the complementarity determining
CC regions of the Ab. These sequences can also be used in a method for
CC determining the number of specific cells in a composition. The methods
CC can be used for positive selection of CD34 haematopoietic cells, and
CC removal of undesired tumour cells or lymphocytes from the selected cells.
CC The pure CD34 cells are reinfused into a cancer patient after high dose
CC therapy to reconstitute the immune system. Cancer cells can also be
CC isolated to determine their chemotherapeutic susceptibility, or for the
CC production of patient-specific vaccines or antibodies. These sequences
CC allow for the recovery of cells from solid supports, in high yield,
CC without enzymatic cleavage. They are relatively inexpensive and safe, and
CC leave cell surface proteins intact
XX
SQ Sequence 6 AA;

Query Match 37.8%; Score 28; DB 2; Length 6;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RISGVG 13
|:||||

```
Db          1 RVSGVG 6


RESULT 6
AAY55175
ID   AAY55175 standard; peptide; 6 AA.
XX
AC   AAY55175;
XX
DT   07-JAN-2000  (first entry)
XX
DE   Anti CD34 antibody 561 releasing peptide SEQ ID NO:69.
XX
KW   Antibody releasing peptide; CD34; hybridoma; binding; antigen;
KW   cell surface antigen; identification; haematopoietic stem cell; tumour;
KW   cancer; immune system; therapy; displacement.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   US5968753-A.
XX
PD   19-OCT-1999.
XX
PF   07-JUN-1995;   95US-00482228.
XX
PR   14-JUN-1994;   94US-00259427.
XX
PA   (NEXE-) NEXELL THERAPEUTICS INC.
XX
PI   Guillermo R,  Helgerson SL,  Deans RJ,  Tseng-Law J,  Kobori JA;
PI   Al-Abdaly FA;
XX
DR   WPI; 1999-590399/50.
XX
PT   Short peptides useful for displacing antibodies from cell surface
PT   antigens.
XX
PS   Claim 7; Col 145; 81pp; English.
XX
CC   The present invention describes peptides of 4-17 amino acids which
CC   displace either the anti-CD34 monoclonal antibody designated 561, the
CC   anti-CD34 mouse monoclonal antibody produced by the hybridoma ATCC HB-
CC   11646 (designated 9069), the anti-CD34 antibody produced by hybridoma
CC   ATCC HB-11885 (9079), or the anti-human breast cancer antibody produced
CC   by hybridoma ATCC HB-11884 (9187), from a cell surface antigen on a
CC   target cell. The peptides are useful for displacing antibodies bound to
CC   cell surfaces to release cells that have been positively selected by
CC   antibody-mediated binding to beads or other solid support. AAY55107 to
CC   AAY55319 represent peptides used in the exemplification of the present
CC   invention
XX
SQ   Sequence 6 AA;

  Query Match             37.8%;  Score 28;  DB 2;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 RISGVG 13
              |:||||
Db          1 RVSGVG 6


RESULT 7
AAY55183
ID   AAY55183 standard; peptide; 6 AA.
XX
AC   AAY55183;
XX
DT   07-JAN-2000  (first entry)
XX
DE   Anti CD34 antibody 561 releasing peptide SEQ ID NO:77.
XX
KW   Antibody releasing peptide; CD34; hybridoma; binding; antigen;
KW   cell surface antigen; identification; haematopoietic stem cell; tumour;
KW   cancer; immune system; therapy; displacement.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   US5968753-A.
XX
PD   19-OCT-1999.
XX
PF   07-JUN-1995;   95US-00482228.
XX
PR   14-JUN-1994;   94US-00259427.
XX
PA   (NEXE-) NEXELL THERAPEUTICS INC.
XX
PI   Guillermo R,  Helgerson SL,  Deans RJ,  Tseng-Law J,  Kobori JA;
PI   Al-Abdaly FA;
XX
DR   WPI; 1999-590399/50.
XX
PT   Short peptides useful for displacing antibodies from cell surface
PT   antigens.
XX
PS   Claim 7; Col 145; 81pp; English.
XX
CC   The present invention describes peptides of 4-17 amino acids which
CC   displace either the anti-CD34 monoclonal antibody designated 561, the
CC   anti-CD34 mouse monoclonal antibody produced by the hybridoma ATCC HB-
CC   11646 (designated 9069), the anti-CD34 antibody produced by hybridoma
CC   ATCC HB-11885 (9079), or the anti-human breast cancer antibody produced
CC   by hybridoma ATCC HB-11884 (9187), from a cell surface antigen on a
CC   target cell. The peptides are useful for displacing antibodies bound to
CC   cell surfaces to release cells that have been positively selected by
CC   antibody-mediated binding to beads or other solid support. AAY55107 to
CC   AAY55319 represent peptides used in the exemplification of the present
CC   invention
XX
SQ   Sequence 6 AA;

  Query Match             37.8%;  Score 28;  DB 2;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 RISGVG 13
              |:||||
Db          1 RVSGVG 6


RESULT 8
AAY86929
ID   AAY86929 standard; peptide; 6 AA.
XX
AC   AAY86929;
XX
DT   09-MAY-2000  (first entry)
XX
DE   Human haematopoietic CD34+ cell binding peptide SEQ ID #77.
XX
KW   Human; haematopoietic CD34+ cell; binding peptide; monoclonal antibody;
KW   non-enzymatic cell selection method; haematopoietic stem cell;
KW   haematopoietic progenitor cell; antibody 561; breast cancer cell;
KW   antibody 9187; cell surface determinant; diagnostic cell based assay.
XX
OS   Homo sapiens.
XX
PN   US6017719-A.
XX
PD   25-JAN-2000.
XX
PF   07-JUN-1995;   95US-00482528.
XX
```

PR 14-JUN-1994; 94US-00259427.
XX
PA (NEXE-) NEXELL THERAPEUTICS INC.
XX
PI Guillermo R, Helgerson SL, Deans RJ, Tseng-Law J, Kobori JA;
PI Al-Abdaly FA;
XX
DR WPI; 2000-136676/12.
XX
PT Non-enzymatic method for the positive selection of target cells from a
PT heterogeneous cell suspension, useful for selecting human breast cancer
PT cells from a patient's blood or bone marrow.
XX
PS Disclosure; Col 19; 82pp; English.
XX
CC This sequence represents a human haematopoietic CD34+ cell binding
CC peptide, and was used to test the method of the invention. The method is
CC a non-enzymatic method for the positive selection of one or more target
CC cells from a heterogeneous cell suspension, by using specific peptides
CC which effect the displacement and release of a specific target cell from
CC a specific monoclonal antibody. The method is useful for positive
CC selection and specific release of target human haematopoietic
CC stem/progenitor cells bound by the monoclonal anti-CD34 antibodies and
CC the antibody 561. The method is also useful for positive selection and
CC specific release of target human breast cancer cells, bound by the
CC monoclonal anti-breast cancer antibody 9187, from a patient's blood or
CC bone marrow. Identification of peptide epitopes for antibodies which
CC recognise cell surface determinants also allows construction of
CC diagnostic cell based assays. The peptide mediated release is enzyme free
CC and thus leaves the cell surface proteins intact. Moreover, peptide
CC mediated release leaves the target cell free of bound antibody or
CC antibody fragments. The method also produces a high yield of functional
CC target cells and is relatively inexpensive to carry out
XX
SQ Sequence 6 AA;

Query Match 37.8%; Score 28; DB 3; Length 6;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RISGVG 13
|:||||
Db 1 RVSGVG 6

RESULT 9
AAY86921
ID AAY86921 standard; peptide; 6 AA.
XX
AC AAY86921;
XX
DT 09-MAY-2000 (first entry)
XX
DE Human haematopoietic CD34+ cell binding peptide SEQ ID #69.
XX
KW Human; haematopoietic CD34+ cell; binding peptide; monoclonal antibody;
KW non-enzymatic cell selection method; haematopoietic stem cell;
KW haematopoietic progenitor cell; antibody 561; breast cancer cell;
KW antibody 9187; cell surface determinant; diagnostic cell based assay.
XX
OS Homo sapiens.
XX
PN US6017719-A.
XX
PD 25-JAN-2000.
XX
PF 07-JUN-1995; 95US-00482528.
XX
PR 14-JUN-1994; 94US-00259427.
XX
PA (NEXE-) NEXELL THERAPEUTICS INC.
XX
PI Guillermo R, Helgerson SL, Deans RJ, Tseng-Law J, Kobori JA;
PI Al-Abdaly FA;
XX
DR WPI; 2000-136676/12.
XX
PT Non-enzymatic method for the positive selection of target cells from a
PT heterogeneous cell suspension, useful for selecting human breast cancer
PT cells from a patient's blood or bone marrow.
XX
PS Example 13; Col 19; 82pp; English.
XX
CC This sequence represents a human haematopoietic CD34+ cell binding
CC peptide, and was used to test the method of the invention. The method is
CC a non-enzymatic method for the positive selection of one or more target
CC cells from a heterogeneous cell suspension, by using specific peptides
CC which effect the displacement and release of a specific target cell from
CC a specific monoclonal antibody. The method is useful for positive
CC selection and specific release of target human haematopoietic
CC stem/progenitor cells bound by the monoclonal anti-CD34 antibodies and
CC the antibody 561. The method is also useful for positive selection and
CC specific release of target human breast cancer cells, bound by the
CC monoclonal anti-breast cancer antibody 9187, from a patient's blood or
CC bone marrow. Identification of peptide epitopes for antibodies which
CC recognise cell surface determinants also allows construction of
CC diagnostic cell based assays. The peptide mediated release is enzyme free
CC and thus leaves the cell surface proteins intact. Moreover, peptide
CC mediated release leaves the target cell free of bound antibody or
CC antibody fragments. The method also produces a high yield of functional
CC target cells and is relatively inexpensive to carry out
XX
SQ Sequence 6 AA;

Query Match 37.8%; Score 28; DB 3; Length 6;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RISGVG 13
|:||||
Db 1 RVSGVG 6

RESULT 10
ADH40843
ID ADH40843 standard; peptide; 9 AA.
XX
AC ADH40843;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human CD64 HLA binding peptide.
XX
KW human; cytostatic; vaccine; SNP profile; cancer; leukaemia; HLA.
XX
OS Homo sapiens.
XX
PN WO2003106692-A2.
XX
PD 24-DEC-2003.
XX
PF 13-JUN-2003; 2003WO-EP006251.
XX
PR 13-JUN-2002; 2002EP-00013423.
XX
PA (MERE) MERCK PATENT GMBH.
XX
PI Strittmatter W, Moll H;
XX
DR WPI; 2004-082200/08.
XX
PT Providing allelic variant epitope of protein based on single nucleotide
PT polymorphism by defining target protein, screening database of protein,
PT identifying, selecting allelic variant protein, creating variant
PT epitopes.
XX

PS Disclosure; Page 98; 119pp; English.
XX
CC The invention relates to a novel method for providing epitopes of allelic
CC variants of antigenic proteins from specific species based on single
CC nucleotide polymorphism (SNP), by defining target protein/peptide or its
CC subset, screening database of DNA encoding target protein, identifying,
CC selecting allelic peptide/protein variants, expression product or its
CC fragment encoded by DNA sequence having SNP, creating variant epitopes,
CC selecting epitopes binding to MHC protein. A protein of the invention has
CC cytostatic activity, and may have a use in a vaccine. The method is
CC useful for generating a SNP profile of one or more individuals from a
CC given species by applying the method for several protein from the
CC individuals, where the SNP profile was related to disease, preferably
CC cancer. This is useful for diagnosing a disease in an individual by
CC generating the SNP-related polymorphic profile. A method of the invention
CC is useful for transplanting haematopoietic stem cells from a donor to a
CC recipient and treating cancer, preferably leukaemia, and for determining
CC the progression, regression or onset of a treated disease. The present
CC sequence is used in the exemplification of the invention.
XX
SQ Sequence 9 AA;

Query Match 37.8%; Score 28; DB 8; Length 9;
Best Local Similarity 62.5%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 TYRISGVG 13
|| ||:|
Db 1 TYHCSGMG 8

RESULT 11
ADN65292
ID ADN65292 standard; peptide; 15 AA.
XX
AC ADN65292;
XX
DT 01-JUL-2004 (first entry)
XX
DE HLA binding peptide #1892.
XX
KW cytostatic; hepatotropic; virucide; antiinflammatory; anti-HIV;
KW gene therapy; vaccine; HLA binding peptide; HTL epitope; liposome;
KW prostate specific antigen; prostate specific membrane antigen;
KW hepatitis B virus antigen; hepatitis C virus antigen;
KW malignant melanoma antigen; MAGE; Epstein Barr virus; cancer;
KW prostate cancer; AIDS; renal carcinoma; cervical carcinoma; lymphoma;
KW chondyloma acuminatum.
XX
OS Unidentified.
XX
PN WO2004031211-A2.
XX
PD 15-APR-2004.
XX
PF 03-OCT-2003; 2003WO-US031308.
XX
PR 03-OCT-2002; 2002US-0416207P.
PR 08-OCT-2002; 2002US-0417269P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sidney J, Southwood S, Sette A;
XX
DR WPI; 2004-347953/32.
XX
PT New composition of peptides and nucleic acids capable of binding Major
PT Histocompatibility Complex molecules, useful for diagnosing, preventing
PT or treating viral infections or cancer, such as prostate cancer,
PT hepatitis B or AIDS.
XX
PS Claim 1; SEQ ID NO 1892; 186pp; English.
XX
CC The invention relates to a novel composition comprising one or more
CC peptides or nucleic acids encoding an HLA binding peptide. The
CC composition further comprises an HTL epitope. It also comprises a spacer
CC molecule, a carrier, an MHC targeting sequence or a lipid. The peptides
CC are incorporated as part of a liposome. The peptide is from an antigen
CC selected from prostate specific antigen (PSA), prostate specific membrane
CC antigen (PSM), hepatitis B virus (HBV) antigen, hepatitis C virus (HCV)
CC antigen, malignant melanoma antigen (MAGE), Epstein Barr virus, human
CC immunodeficiency type-1 (HIV-I), human immunodeficiency type-2 (HIV-2),
CC Papilloma virus, Lassa virus, Mycobacterium tuberculosis (MT), p53,
CC murine p53 (mp53), CEA, HER2/neu, and tyrosine kinase related protein
CC (TKP). The composition is useful for preventing or treating viral
CC infections or cancer, such as prostate cancer, hepatitis B, hepatitis C,
CC AIDS, renal carcinoma, cervical carcinoma, lymphoma, CMV or chondyloma
CC acuminatum. The composition is also be used for diagnosing such diseases.
CC This sequence represents a peptide of the invention.
XX
SQ Sequence 15 AA;

Query Match 37.8%; Score 28; DB 8; Length 15;
Best Local Similarity 40.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TQKITYRISG 11
|||: ::::|
Db 2 TQKLLWQLNG 11

RESULT 12
AAG86324
ID AAG86324 standard; peptide; 10 AA.
XX
AC AAG86324;
XX
DT 11-SEP-2001 (first entry)
XX
DE Saccharomyces cerevisiae peptide, SEQ ID NO: 1273.
XX
KW Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW drug discovery; drug design.
XX
OS Saccharomyces cerevisiae.
XX
PN WO200142276-A1.
XX
PD 14-JUN-2001.
XX
PF 13-DEC-2000; 2000WO-GB004773.
XX
PR 13-DEC-1999; 99GB-00029471.
XX
PA (PROT-) PROTEOM LTD.
XX
PI Roberts GW, Heal JR;
XX
DR WPI; 2001-367863/38.
XX
PT Identifying complementary peptides by analysis of protein and nucleotide
PT sequence databases, useful in drug design.
XX
PS Example 3; Page 205; 488pp; English.
XX
CC The invention relates to the identification of complementary peptides by
CC analysis of protein and nucleotide sequence databases from higher
CC eukaryotic genomes, excluding human and plants. The specific
CC complementary peptides interact with their relevant target proteins
CC encoded in the eukaryote genome. The peptides may be used as reagents and
CC drugs for drug discovery and as lead ligands for drug design and
CC development. The present sequence is a complementary peptide from
CC Saccharomyces cerevisiae
XX
SQ Sequence 10 AA;

```
  Query Match             36.5%;  Score 27;  DB 4;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 4.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 ISGVGI 14
              |||||:
Db          5 ISGVGV 10


RESULT 13
AAG86188
ID   AAG86188 standard; peptide; 10 AA.
XX
AC   AAG86188;
XX
DT   11-SEP-2001  (first entry)
XX
DE   Saccharomyces cerevisiae peptide, SEQ ID NO: 1137.
XX
KW   Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW   drug discovery; drug design.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO200142276-A1.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB004773.
XX
PR   13-DEC-1999;   99GB-00029471.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-367863/38.
XX
PT   Identifying complementary peptides by analysis of protein and nucleotide
PT   sequence databases, useful in drug design.
XX
PS   Example 3; Page 186; 488pp; English.
XX
CC   The invention relates to the identification of complementary peptides by
CC   analysis of protein and nucleotide sequence databases from higher
CC   eukaryotic genomes, excluding human and plants. The specific
CC   complementary peptides interact with their relevant target proteins
CC   encoded in the eukaryote genome. The peptides may be used as reagents and
CC   drugs for drug discovery and as lead ligands for drug design and
CC   development. The present sequence is a complementary peptide from
CC   Saccharomyces cerevisiae
XX
SQ   Sequence 10 AA;

  Query Match             36.5%;  Score 27;  DB 4;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 4.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 ISGVGI 14
              |||||:
Db          5 ISGVGV 10


RESULT 14
ADN70236
ID   ADN70236 standard; peptide; 10 AA.
XX
AC   ADN70236;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human 273P4B7v.1 protein epitope #4199.
XX
KW   273P4B7; human; cancer; tumour; epitope.
XX
OS   Homo sapiens.
XX
PN   WO2004016762-A2.
XX
PD   26-FEB-2004.
XX
PF   15-AUG-2003; 2003WO-US025665.
XX
PR   16-AUG-2002; 2002US-0404306P.
PR   01-NOV-2002; 2002US-0423290P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Challita-Eid PM,  Faris M,  Raitano AB,  Jakobovits A,  Ge W;
XX
DR   WPI; 2004-203790/19.
XX
PT   New composition comprising 273P4B7 proteins, useful for detecting and
PT   treating cancer by inhibiting the growth or viability of cancer cells.
XX
PS   Claim 1; Fig 2A; 268pp; English.
XX
CC   The invention relates to a composition comprising 273P4B7 proteins. The
CC   composition and proteins are useful for detecting and treating cancer by
CC   inhibiting the growth or viability of cancer cells. The present sequence
CC   represents the amino acid sequence of a human 273P4B7v.1 protein epitope.
CC   Note the epitope sequences are displayed in tables VIII-XLIX.
XX
SQ   Sequence 10 AA;

  Query Match             36.5%;  Score 27;  DB 8;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 4.3e+02;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 ITYRISGVGI 14
              :| :: |||:
Db          1 LTTQVGGVGL 10


RESULT 15
ADN67283
ID   ADN67283 standard; peptide; 10 AA.
XX
AC   ADN67283;
XX
DT   01-JUL-2004  (first entry)
XX
DE   Human 273P4B7v.1 protein epitope #1246.
XX
KW   273P4B7; human; cancer; tumour; epitope.
XX
OS   Homo sapiens.
XX
PN   WO2004016762-A2.
XX
PD   26-FEB-2004.
XX
PF   15-AUG-2003; 2003WO-US025665.
XX
PR   16-AUG-2002; 2002US-0404306P.
PR   01-NOV-2002; 2002US-0423290P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Challita-Eid PM,  Faris M,  Raitano AB,  Jakobovits A,  Ge W;
XX
DR   WPI; 2004-203790/19.
XX
PT   New composition comprising 273P4B7 proteins, useful for detecting and
PT   treating cancer by inhibiting the growth or viability of cancer cells.
XX
```

PS Claim 1; Fig 2A; 268pp; English.
XX
CC The invention relates to a composition comprising 273P4B7 proteins. The
CC composition and proteins are useful for detecting and treating cancer by
CC inhibiting the growth or viability of cancer cells. The present sequence
CC represents the amino acid sequence of a human 273P4B7v.1 protein epitope.
CC Note the epitope sequences are displayed in tables VIII-XLIX.
XX
SQ Sequence 10 AA;

Query Match 36.5%; Score 27; DB 8; Length 10;
Best Local Similarity 40.0%; Pred. No. 4.3e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```
Qy          5 ITYRISGVGI 14
              :| :: |||:
Db          1 LTTQVGGVGL 10
```

Search completed: February 22, 2005, 09:24:32
Job time : 68.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-2
Perfect score: 78
Sequence: 1 FGIFVVDKNTGDINI 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	30.8	15	2	PA0052	protein QF200015 -
2	23	29.5	10	2	A49581	sialokinin I - yel
3	23	29.5	11	2	A40795	glycoprotein H-a -
4	23	29.5	14	2	B58502	36K kidney stone p
5	23	29.5	15	2	PQ0017	terminal protein -
6	22	28.2	14	2	B44854	L-2,4-diaminobutyr
7	22	28.2	15	2	A47146	topoisomerase I -
8	22	28.2	15	2	S66215	cartilage oligomer
9	20	25.6	13	2	A26999	carboxylesterase (
10	20	25.6	14	2	S09721	2S albumin small c
11	19	24.4	11	2	S51732	T-cell receptor al
12	19	24.4	11	2	PH0919	T-cell receptor be
13	19	24.4	11	4	PC2124	aminotransferase c
14	19	24.4	12	2	S36899	ribosomal protein
15	19	24.4	12	2	S71034	potB protein - Sal
16	19	24.4	14	2	PH1614	Ig H chain V-D-J r
17	19	24.4	14	2	PH1617	Ig H chain V-D-J r
18	19	24.4	14	2	A44515	Trp EG leader pept
19	19	24.4	15	2	PH0750	T-cell receptor be
20	18	23.1	10	2	S10926	inhibin beta-A cha
21	18	23.1	10	2	B49581	sialokinin II - ye
22	18	23.1	11	2	S23926	major glycoprotein
23	18	23.1	13	2	PH0796	T-cell receptor al
24	18	23.1	13	2	PH0799	T-cell receptor al
25	18	23.1	13	2	PH0783	T-cell receptor al
26	18	23.1	14	2	S22236	lipoxygenase (EC 1
27	18	23.1	14	2	S60353	amylopullulanase -
28	18	23.1	15	2	S43956	Ig mu chain V regi
29	18	23.1	15	2	S47367	T-cell antigen rec
30	18	23.1	15	2	S55312	TSH protein beta c
31	18	23.1	15	2	A35141	T-cell receptor de
32	18	23.1	15	2	PC4213	bphB protein - Com
33	17.5	22.4	15	2	PH1320	Ig heavy chain DJ
34	17	21.8	6	2	A61419	sarcosine dehydrog
35	17	21.8	7	2	S45311	microcin C7 - Esch
36	17	21.8	11	2	PT0211	T-cell receptor al
37	17	21.8	11	2	S45386	low density lipopr
38	17	21.8	12	2	G49410	t-complex polypept
39	17	21.8	12	2	C39109	hypothetical 1.2K
40	17	21.8	12	2	PT0257	Ig heavy chain CRD
41	17	21.8	12	2	PT0216	T-cell receptor be
42	17	21.8	12	2	PH0920	T-cell receptor be
43	17	21.8	12	2	PN0160	ribosomal protein
44	17	21.8	13	1	JTJG3	tremerogen a-13 -
45	17	21.8	13	2	S29488	GTP-binding protei

ALIGNMENTS

RESULT 1
PA0052
protein QF200015 - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C;Accession: PA0052
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0052
A;Molecule type: protein
A;Residues: 1-15 <CHO>

Query Match 30.8%; Score 24; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy          3 IFVVDKNTGDIN 14
              || ||    | |
Db          3 IFXVDAVYNDTN 14
```

RESULT 2
A49581
sialokinin I - yellow fever mosquito
C;Species: Aedes aegypti (yellow fever mosquito)
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A49581
R;Champagne, D.E.; Ribeiro, J.M.
Proc. Natl. Acad. Sci. U.S.A. 91, 138-142, 1994
A;Title: Sialokinin I and II: vasodilatory tachykinins from the yellow fever mosquito A
A;Reference number: A49581; MUID:94105119; PMID:8278354
A;Accession: A49581
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <CHA>
A;Cross-references: UNIPROT:P42634
A;Experimental source: Rockefeller, salivary gland
A;Note: sequence extracted from NCBI backbone (NCBIP:141841)
C;Superfamily: unassigned animal peptides

Query Match 29.5%; Score 23; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 8.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          9 NTGD 12
              ||||
Db          1 NTGD 4
```

RESULT 3

A40795
glycoprotein H-a - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 31-Dec-1993
C;Accession: A40795
R;Christie, D.L.; Batchelor, D.C.; Palmer, D.J.
J. Biol. Chem. 266, 15679-15683, 1991
A;Title: Identification of kex2-related proteases in chromaffin granules by partial amin
A;Reference number: A40795; MUID:91340701; PMID:1874725
A;Accession: A40795
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <CHR>
C;Keywords: glycoprotein

Query Match 29.5%; Score 23; DB 2; Length 11;
Best Local Similarity 55.6%; Pred. No. 9.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 7 DKNTGDINI 15
| | |||:
Db 3 DINEIDINV 11

RESULT 4
B58502
36K kidney stone protein - unidentified bacterium (fragment)
C;Species: unidentified bacterium
C;Date: 07-Feb-1997 #sequence_revision 07-Feb-1997 #text_change 10-Jul-1998
C;Accession: B58502
R;Binette, J.P.; Binette, M.B.
submitted to the Protein Sequence Database, October 1996
A;Description: The proteins of kidney and gallbladder stones.
A;Reference number: A58501
A;Accession: B58502
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <BIN>
A;Experimental source: human kidney stone containing Ca ox.mono& dihyd, 1% struvite, CaP
A;Note: tentative identification of 8-Tyr and 9-Thr

Query Match 29.5%; Score 23; DB 2; Length 14;
Best Local Similarity 71.4%; Pred. No. 1.2e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FGIFVVD 7
|| |||
Db 6 FGYTVVD 12

RESULT 5
PQ0017
terminal protein - phage M2 (fragment)
C;Species: phage M2
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: PQ0017
R;Matsumoto, K.; Takano, H.; Kim, C.I.; Hirokawa, H.
Gene 84, 247-255, 1989
A;Title: Primary structure of bacteriophage M2 DNA polymerase: conserved segments within
A;Reference number: JQ0161; MUID:90128268; PMID:2515115
A;Accession: PQ0017
A;Molecule type: DNA
A;Residues: 1-15 <MAT>
A;Cross-references: UNIPROT:P19897; GB:M33144; NID:g215507; PIDN:AAA32367.1; PID:g215508
C;Genetics:
A;Gene: E
C;Superfamily: phage PZA terminal protein

Query Match 29.5%; Score 23; DB 2; Length 15;
Best Local Similarity 60.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 11 GDINI 15
||:|:
Db 6 GDVNL 10

RESULT 6
B44854
L-2,4-diaminobutyrate decarboxylase (EC 4.1.1.-) - Vibrio alginolyticus (fragment)
C;Species: Vibrio alginolyticus
C;Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B44854; B41817
R;Yamamoto, S.; Tsuzaki, Y.; Tougou, K.; Shinoda, S.
J. Gen. Microbiol. 138, 1461-1465, 1992
A;Title: Purification and characterization of L-2,4-diaminobutyrate decarboxylase from A
A;Reference number: A44854; MUID:92381494; PMID:1512577
A;Accession: B44854
A;Molecule type: protein
A;Residues: 1-14 <YAM>
A;Cross-references: UNIPROT:Q9R5I8
A;Note: sequence extracted from NCBI backbone (NCBIP:112332)
C;Keywords: carbon-carbon lyase; carboxy-lyase

Query Match 28.2%; Score 22; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 4 FVVDKNTGDI 13
| || | :|
Db 4 FEVDSNIWNI 13

RESULT 7
A47146
topoisomerase I - vaccinia virus (fragment)
C;Species: vaccinia virus
C;Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 09-Jul-2004
C;Accession: A47146
R;Klemperer, N.; Traktman, P.
J. Biol. Chem. 268, 15887-15899, 1993
A;Title: Biochemical analysis of mutant alleles of the vaccinia virus topoisomerase I ca
A;Reference number: A47146; MUID:93340198; PMID:8393454
A;Accession: A47146
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-15 <KLE>
A;Cross-references: UNIPROT:Q9JFB0; GB:L13447
C;Superfamily: vaccinia virus DNA topoisomerase

Query Match 28.2%; Score 22; DB 2; Length 15;
Best Local Similarity 38.5%; Pred. No. 2e+03;
Matches 5; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 2 GIFVVDKNTGDIN 14
|| : | | :|
Db 1 GIRIKDLRTYGVN 13

RESULT 8
S66215
cartilage oligomeric matrix protein - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C;Accession: S66215
R;Hauser, N.; Paulsson, M.; Kale, A.A.; DiCesare, P.E.
FEBS Lett. 368, 307-310, 1995
A;Title: Tendon extracellular matrix contains pentameric thrombospondin-4 (TSP-4).
A;Reference number: S66214; MUID:95354859; PMID:7628627
A;Accession: S66215
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <HAU>

Query Match 28.2%; Score 22; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 2e+03;

Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          9 NTGDIN 14
              | ||| |
Db          8 NGGDFN 13
```

RESULT 9
A26999
carboxylesterase (EC 3.1.1.1), intestinal - Caenorhabditis elegans (fragment)
C;Species: Caenorhabditis elegans
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 09-Jul-2004
C;Accession: A26999
R;McGhee, J.D.
Biochemistry 26, 4101-4107, 1987
A;Title: Purification and characterization of a carboxylesterase from the intestine of t
A;Reference number: A26999; MUID:88000636; PMID:3651439
A;Accession: A26999
A;Molecule type: protein
A;Residues: 1-13 <MCG>
A;Cross-references: UNIPROT:Q7M3Q8
C;Keywords: carboxylic ester hydrolase; intestine

Query Match 25.6%; Score 20; DB 2; Length 13;
Best Local Similarity 33.3%; Pred. No. 3.6e+03;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy          5 VVDKNTGDI 13
              :|: | | :
Db          3 IVEHNYGKV 11
```

RESULT 10
S09721
2S albumin small chain nIII - rape (fragments)
C;Species: Brassica napus (rape)
C;Date: 19-Mar-1997 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998
C;Accession: S09721
R;Monsalve, R.I.; Menendez-Arias, L.; Lopez-Otin, C.; Rodriguez, R.
FEBS Lett. 263, 209-212, 1990
A;Title: beta-Turns as structural motifs for the proteolytic processing of seed proteins
A;Reference number: S09720; MUID:90242974; PMID:2185951
A;Accession: S09721
A;Molecule type: protein
A;Residues: 1-9;10-14 <MON>
A;Experimental source: seed

Query Match 25.6%; Score 20; DB 2; Length 14;
Best Local Similarity 40.0%; Pred. No. 3.9e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 GIFVVDKNTG 11
              | | : | :|
Db          3 GPFRIPKQSG 12
```

RESULT 11
S51732
T-cell receptor alpha chain joining region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 05-Nov-1999
C;Accession: S51732
R;Durinovic-Bello, I.; Steinle, A.; Ziegler, A.G.; Schendel, D.J.
submitted to the EMBL Data Library, November 1993
A;Reference number: S51732
A;Accession: S51732
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-11 <DUR>
A;Cross-references: EMBL:Z28343; NID:g607116; PIDN:CAA82197.1; PID:g607117
C;Keywords: T-cell receptor

Query Match 24.4%; Score 19; DB 2; Length 11;
Best Local Similarity 50.0%; Pred. No. 4.4e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          8 KNTGDI 13
              :||| :
Db          4 ENTGKL 9
```

RESULT 12
PH0919
T-cell receptor beta chain V-D-J region (isolate 5) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C;Accession: PH0919
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergi
A;Reference number: PH0891; MUID:92078857; PMID:1836012
A;Accession: PH0919
A;Molecule type: mRNA
A;Residues: 1-11 <GOL>
A;Experimental source: concanavalin A-activated lymphoblast
A;Note: the authors translated the codon CAG for residue 11 as Glu
C;Keywords: T-cell receptor

Query Match 24.4%; Score 19; DB 2; Length 11;
Best Local Similarity 75.0%; Pred. No. 4.4e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          8 KNTG 11
              :|||
Db          7 RNTG 10
```

RESULT 13
PC2124
aminotransferase chimera DY376 - synthetic (fragment)
C;Species: synthetic
C;Date: 28-May-1999 #sequence_revision 28-May-1999 #text_change 28-May-1999
C;Accession: PC2124
R;Miyazawa, K.; Kawaguchi, S.; Okamoto, A.; Kato, R.; Ogawa, T.; Kuramitsu, S.
J. Biochem. 115, 568-577, 1994
A;Title: Construction of aminotransferase chimeras and analysis of their substrate speci
A;Reference number: JX0315; MUID:94334304; PMID:8056774
A;Accession: PC2124
A;Molecule type: DNA
A;Residues: 1-11 <MIY>
C;Comment: This is a chimeric enzyme of Escherichia coli aspartate aminotransferase (EC
C;Comment: The parental enzymes catalyze the reversible amino group transfer reaction bc
C;Genetics:
A;Gene: aspC; tyrB
C;Keywords: aminotransferase

Query Match 24.4%; Score 19; DB 4; Length 11;
Best Local Similarity 40.0%; Pred. No. 4.4e+03;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 FGIFV 5
              ||:::
Db          7 FGVYL 11
```

RESULT 14
S36899
ribosomal protein S6 - Mycobacterium bovis (fragment)
C;Species: Mycobacterium bovis
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 13-Jan-1995
C;Accession: S36899
R;Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.
FEBS Lett. 331, 9-14, 1993
A;Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Mycoba
A;Reference number: S36887; MUID:94009653; PMID:8405418
A;Accession: S36899

```
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <OHA>
C;Keywords: protein biosynthesis; ribosome

  Query Match             24.4%;  Score 19;  DB 2;  Length 12;
  Best Local Similarity   57.1%;  Pred. No. 4.8e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 FGIFVVD 7
              : | |||
Db          4 YEIMVVD 10


RESULT 15
S71034
potB protein - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S71034
R;Stein, M.A.; Leung, K.Y.; Zwick, M.; Garcia-del Portillo, F.; Finlay, B.B.
Mol. Microbiol. 20, 151-164, 1996
A;Title: Identification of a Salmonella virulence gene required for formation of filamen
A;Reference number: S71033; MUID:97014378; PMID:8861213
A;Accession: S71034
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-12 <STE>
A;Cross-references: UNIPROT:Q56060; EMBL:U51867; NID:g1272352; PIDN:AAA97466.1; PID:g127
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1996
C;Genetics:
A;Gene: potB

  Query Match             24.4%;  Score 19;  DB 2;  Length 12;
  Best Local Similarity   30.0%;  Pred. No. 4.8e+03;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 VVDKNTGDIN 14
              :::|   ||:
Db          2 LLNKKVSDIS 11


Search completed: February 22, 2005, 09:46:23
Job time : 12.1333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-2
Perfect score: 78
Sequence: 1 FGIFVVDKNTGDINI 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	34.6	12	2	Q8J1G6	Q8j1g6 ashbya goss
2	25	32.1	13	2	P97140	P97140 borrelia bu
3	25	32.1	15	2	Q9R4P6	Q9r4p6 brevundimon
4	24	30.8	13	2	Q8WY56	Q8wy56 homo sapien
5	23	29.5	10	1	TKS1_AEDAE	P42634 aedes aegyp
6	23	29.5	12	2	Q9TQY5	Q9tgy5 bos taurus
7	23	29.5	14	2	Q9S8X6	Q9s8x6 glycine max
8	23	29.5	14	2	Q65CH4	Q65ch4 sinaloa tom
9	23	29.5	15	1	TERM_BPM2	P19897 bacteriopha
10	22	28.2	11	2	Q997C1	Q997c1 east africa
11	22	28.2	14	2	Q9R5I8	Q9r5i8 vibrio algi
12	22	28.2	15	2	Q75LP0	Q75lp0 homo sapien
13	21	26.9	12	2	Q6SJX6	Q6sjx6 photobacter
14	21	26.9	13	2	Q9UPE5	Q9upe5 homo sapien
15	21	26.9	13	2	Q6LBZ3	Q6lbz3 vibrio harv
16	21	26.9	13	2	Q6VFN6	Q6vfn6 photobacter
17	21	26.9	14	1	UC18_MAIZE	P80624 zea mays (m
18	21	26.9	14	1	UN46_CLOPA	P81362 clostridium
19	21	26.9	14	2	Q9RPG9	Q9rpg9 listeria mo
20	21	26.9	15	2	Q7S362	Q7s362 neurospora
21	21	26.9	15	2	Q9S8D4	Q9s8d4 cynara card
22	20	25.6	9	1	UPA6_HUMAN	P30092 homo sapien
23	20	25.6	11	2	P83128	P83128 bos indicus
24	20	25.6	12	1	PA2B_VIPBO	P31859 vipera beru
25	20	25.6	13	2	Q7M3Q8	Q7m3q8 caenorhabdi
26	20	25.6	13	2	Q9THS3	Q9ths3 bryopsis sp
27	20	25.6	15	2	Q9R597	Q9r597 micrococcus
28	20	25.6	15	2	Q9RFZ5	Q9rfz5 mycoplasma
29	19	24.4	10	2	Q9S936	Q9s936 beta vulgar
30	19	24.4	13	2	Q53693	Q53693 streptomyce
31	19	24.4	14	1	LPW_RHIME	P18854 rhizobium m
32	19	24.4	14	2	Q9G3C0	Q9g3c0 calliphora
33	19	24.4	15	1	UC28_MAIZE	P80634 zea mays (m
34	19	24.4	15	2	Q69FJ1	Q69fj1 stemphylium
35	19	24.4	15	2	Q86TJ6	Q86tj6 homo sapien
36	19	24.4	15	2	Q9TR52	Q9tr52 bos taurus
37	19	24.4	15	2	Q9S8I1	Q9s8i1 volvox cart
38	18	23.1	8	2	Q6SE42	Q6se42 drosophila
39	18	23.1	9	2	Q70Y81	Q70y81 plectranthu
40	18	23.1	10	1	TKS2_AEDAE	P42635 aedes aegyp
41	18	23.1	11	2	Q9UAR8	Q9uar8 aedes aegyp
42	18	23.1	11	2	Q7M2M1	Q7m2m1 bos taurus
43	18	23.1	11	2	O77910	O77910 oreochromis
44	18	23.1	12	1	HCYB_MEGCR	Q10584 megathura c
45	18	23.1	13	2	Q69GD7	Q69gd7 pleospora t

ALIGNMENTS

RESULT 1
Q8J1G6
ID Q8J1G6 PRELIMINARY; PRT; 12 AA.
AC Q8J1G6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE YEL062 (Fragment).
GN Name=YEL062;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RA Alberti-Segui C., Dietrich F., Philippsen P.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF378568; AAN87135.1; -.
FT NON_TER 1 1
SQ SEQUENCE 12 AA; 1294 MW; 2EE79139285B5B18 CRC64;

Query Match 34.6%; Score 27; DB 2; Length 12;
Best Local Similarity 62.5%; Pred. No. 1.2e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          8 KNTGDINI 15
              | ||: ||
Db          2 KGTGEYNI 9
```

RESULT 2
P97140
ID P97140 PRELIMINARY; PRT; 13 AA.
AC P97140;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE Flagellin protein (Fragment).
GN Name=flaB;
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=212;
RX MEDLINE=97312006; PubMed=9168617;
RA Ge Y., Old I.G., Girons I.S., Charon N.W.;
RT "The flgK motility operon of Borrelia burgdorferi is initiated by a
RT sigma 70-like promoter.";
RL Microbiology 143:1681-1690(1997).
DR EMBL; U66699; AAB58987.1; -.
DR GO; GO:0019861; C:flagellum; IEA.
KW Flagellum.
FT NON_TER 13 13

```
SQ   SEQUENCE   13 AA;  1386 MW;  188978372FFCA1A4 CRC64;

  Query Match             32.1%;  Score 25;  DB 2;  Length 13;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+03;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 VVDKNTGDIN 14
              ::: ||  ||
Db          2 IINHNTSAIN 11


RESULT 3
Q9R4P6
ID   Q9R4P6      PRELIMINARY;      PRT;    15 AA.
AC   Q9R4P6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   Ribosomal protein S21 (Fragment).
OS   Brevundimonas diminuta (Pseudomonas diminuta).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Brevundimonas.
OX   NCBI_TaxID=293;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95244309; PubMed=7727274;
RA   Ochi K.;
RT   "Comparative ribosomal protein sequence analyses of a phylogenetically
RT   defined genus, Pseudomonas, and its relatives.";
RL   Int. J. Syst. Bacteriol. 45:268-273(1995).
SQ   SEQUENCE   15 AA;  1713 MW;  012C571E5EDBD3CA CRC64;

  Query Match             32.1%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 3.3e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 IFVVDKN 9
              ||| | |
Db          3 IFVXDNN 9


RESULT 4
Q8WY56
ID   Q8WY56      PRELIMINARY;      PRT;    13 AA.
AC   Q8WY56;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Epithelial sodium channel beta-3 subunit (Fragment).
GN   Name=SCNN1B;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=11934701;
RA   Thomas C.P., Loftus R.W., Liu K.Z., Itani O.A.;
RT   "Genomic organization of the 5' end of human beta-ENaC and preliminary
RT   characterization of its promoter.";
RL   Am. J. Physiol. Renal Physiol. 282:F898-F909(2002).
DR   EMBL; AF260228; AAL48197.1; -.
DR   GO; GO:0005216; F:ion channel activity; IEA.
KW   Ionic channel.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1437 MW;  1716D00275917724 CRC64;

  Query Match             30.8%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 4.2e+03;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 DKNTGDIN 14
              | | || |
Db          4 DGNLGDKN 11


RESULT 5
TKS1_AEDAE
ID   TKS1_AEDAE     STANDARD;      PRT;    10 AA.
AC   P42634;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Sialokinin I.
OS   Aedes aegypti (Yellowfever mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
OX   NCBI_TaxID=7159;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=Rockefeller; TISSUE=Salivary gland;
RX   MEDLINE=94105119; PubMed=8278354;
RA   Champagne D.E., Ribeiro J.M.C.;
RT   "Sialokinin I and II: vasodilatory tachykinins from the yellow fever
RT   mosquito Aedes aegypti.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:138-142(1994).
CC   -!- FUNCTION: Vasodilatory peptide. May activate macrophages at the
CC       site of feeding.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the tachykinin family.
DR   PIR; A49581; A49581.
DR   InterPro; IPR002040; Tachy_Neurokinin.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Amidation; Direct protein sequencing; Neuropeptide; Tachykinin.
FT   MOD_RES      10     10       Methionine amide.
SQ   SEQUENCE   10 AA;  1145 MW;  3DCFDE6B59C33AA8 CRC64;

  Query Match             29.5%;  Score 23;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 4.6e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 NTGD 12
              ||||
Db          1 NTGD 4


RESULT 6
Q9TQY5
ID   Q9TQY5      PRELIMINARY;      PRT;    12 AA.
AC   Q9TQY5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Glycoprotein H-A N-TERMINAL, GPH-A N-TERMINAL=KEX2/subtilisin-related
DE   protease (Fragment).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=91340701; PubMed=1874725;
RA   Christie D.L., Batchelor D.C., Palmer D.J.;
RT   "Identification of kex2-related proteases in chromaffin granules by
RT   partial amino acid sequence analysis.";
RL   J. Biol. Chem. 266:15679-15683(1991).
FT   NON_TER       1      1
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1374 MW;  0BDF36703B5B1440 CRC64;

  Query Match             29.5%;  Score 23;  DB 2;  Length 12;
  Best Local Similarity   55.6%;  Pred. No. 5.6e+03;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 DKNTGDINI 15
```

```
              | |  |||:
Db          4 DINEIDINV 12


RESULT 7
Q9S8X6
ID   Q9S8X6      PRELIMINARY;      PRT;    14 AA.
AC   Q9S8X6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Vegetative storage protein 94 peptide 1, VSP94=LIPOXYGENASE
DE   (Fragment).
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92361246; PubMed=1822994;
RA   Tranbarger T.J., Franceschi V.R., Hildebrand D.F., Grimes H.D.;
RT   "The soybean 94-kilodalton vegetative storage protein is a
RT   lipoxygenase that is localized in paraveinal mesophyll cell
RT   vacuoles.";
RL   Plant Cell 3:973-987(1991).
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1541 MW;  98EB730EA6AE785A CRC64;

  Query Match             29.5%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 6.7e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 KNTGDIN 14
              ||  |||
Db          3 KNVLDIN 9


RESULT 8
Q65CH4
ID   Q65CH4      PRELIMINARY;      PRT;    14 AA.
AC   Q65CH4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Coat protein (Fragment).
GN   Name=AV1;
OS   Sinaloa tomato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=71186;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NI1;
RA   Rojas A., Kvarnheden A., Rodriguez D., Valkonen J.P.T.;
RT   "A mixture of begomoviruses in severe leaf curl-affected tomatoes in
RT   Nicaragua.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ508778; CAD48516.1; -.
KW   Coat protein.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1573 MW;  571284313A0594D3 CRC64;

  Query Match             29.5%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   55.6%;  Pred. No. 6.7e+03;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 VDKNTGDIN 14
              ||| |  :|
Db          5 VDKATAWVN 13


RESULT 9
TERM_BPM2
ID   TERM_BPM2      STANDARD;      PRT;    15 AA.
AC   P19897;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   DNA terminal protein (Protein GP3) (Fragment).
GN   Name=3; Synonyms=E;
OS   Bacteriophage M2.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC   phi-29-like viruses.
OX   NCBI_TaxID=10751;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90128268; PubMed=2515115; DOI=10.1016/0378-1119(89)90498-8;
RA   Matsumoto K., Takano H., Kim C.I., Hirokawa H.;
RT   "Primary structure of bacteriophage M2 DNA polymerase: conserved
RT   segments within protein-priming DNA polymerases and DNA polymerase I
RT   of Escherichia coli.";
RL   Gene 84:247-255(1989).
CC   -!- FUNCTION: DNA terminal protein is linked to the 5' ends of both
CC       strands of the genome through a phosphodiester bond between the
CC       beta-hydroxyl group of a serine residue and the 5'-phosphate of
CC       the terminal deoxyadenylate. This protein is essential for DNA
CC       replication and is involved in the priming of DNA elongation.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M33144; AAA32367.1; -.
DR   PIR; PQ0017; PQ0017.
DR   InterPro; IPR008770; Phi-29_GP3.
DR   Pfam; PF05435; Phi-29_GP3; 1.
KW   Covalent protein-DNA linkage; DNA priming; DNA replication;
KW   Early protein.
FT   NON_TER       1      1
FT   SITE          5      7       Cell attachment site (Potential).
SQ   SEQUENCE   15 AA;  1797 MW;  D3CBAFF8759DEA06 CRC64;

  Query Match             29.5%;  Score 23;  DB 1;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 7.2e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GDINI 15
              ||:|:
Db          6 GDVNL 10


RESULT 10
Q997C1
ID   Q997C1      PRELIMINARY;      PRT;    11 AA.
AC   Q997C1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Coat protein (Fragment).
OS   East African cassava mosaic virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=62079;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21103006; PubMed=11172108;
RA   Pita J.S., Fondong V.N., Sangare A., Otim-Nape G.W., Ogwal S.,
RA   Fauquet C.M.;
RT   "Recombination, pseudorecombination and synergism of geminiviruses are
RT   determinant keys to the epidemic of severe cassava mosaic disease in
RT   Uganda.";
RL   J. Gen. Virol. 82:655-665(2001).
```

```
DR   EMBL; AF230374; AAK26738.1; -.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
KW   Coat protein.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1216 MW;  7751D0695AA86774 CRC64;

  Query Match             28.2%;  Score 22;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 7.6e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 VDKNTGDINI 15
              : |  ||| |
Db          1 MSKRPGDIII 10


RESULT 11
Q9R5I8
ID   Q9R5I8      PRELIMINARY;      PRT;    14 AA.
AC   Q9R5I8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   L-2,4-diaminobutyrate decarboxylase (Fragment).
OS   Vibrio alginolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=663;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92381494; PubMed=1512577;
RA   Yamamoto S., Tsuzaki Y., Tougou K., Shinoda S.;
RT   "Purification and characterization of L-2,4-diaminobutyrate
RT   decarboxylase from Acinetobacter calcoaceticus.";
RL   J. Gen. Microbiol. 138:1461-1465(1992).
DR   PIR; B44854; B44854.
SQ   SEQUENCE   14 AA;  1643 MW;  9F1B13DD35168ABA CRC64;

  Query Match             28.2%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 9.9e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 FVVDKNTGDI 13
              | ||| |  :|
Db          4 FEVDSNIWNI 13


RESULT 12
Q75LP0
ID   Q75LP0      PRELIMINARY;      PRT;    15 AA.
AC   Q75LP0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein ARP3BETA (Fragment).
GN   Name=ARP3BETA;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22737999; PubMed=12853948; DOI=10.1038/nature01782;
RA   Hillier L.W., Fulton R.S., Fulton L.A., Graves T.A., Pepin K.H.,
RA   Wagner-McPherson C., Layman D., Maas J., Jaeger S., Walker R.,
RA   Wylie K., Sekhon M., Becker M.C., O'Laughlin M.D., Schaller M.E.,
RA   Fewell G.A., Delehaunty K.D., Miner T.L., Nash W.E., Cordes M., Du H.,
RA   Sun H., Edwards J., Bradshaw-Cordum H., Ali J., Andrews S., Isak A.,
RA   Vanbrunt A., Nguyen C., Du F., Lamar B., Courtney L., Kalicki J.,
RA   Ozersky P., Bielicki L., Scott K., Holmes A., Harkins R., Harris A.,
RA   Strong C.M., Hou S., Tomlinson C., Dauphin-Kohlberg S.,
RA   Kozlowicz-Reilly A., Leonard S., Rohlfing T., Rock S.M.,
RA   Tin-Wollam A.M., Abbott A., Minx P., Maupin R., Strowmatt C.,
RA   Latreille P., Miller N., Johnson D., Murray J., Woessner J.P.,
RA   Wendl M.C., Yang S.P., Schultz B.R., Wallis J.W., Spieth J.,
RA   Bieri T.A., Nelson J.O., Berkowicz N., Wohldmann P.E., Cook L.L.,
RA   Hickenbotham M.T., Eldred J., Williams D., Bedell J.A., Mardis E.R.,
RA   Clifton S.W., Chissoe S.L., Marra M.A., Raymond C., Haugen E.,
RA   Gillett W., Zhou Y., James R., Phelps K., Iadanoto S., Bubb K.,
RA   Simms E., Levy R., Clendenning J., Kaul R., Kent W.J., Furey T.S.,
RA   Baertsch R.A., Brent M.R., Keibler E., Flicek P., Bork P., Suyama M.,
RA   Bailey J.A., Portnoy M.E., Torrents D., Chinwalla A.T., Gish W.R.,
RA   Eddy S.R., McPherson J.D., Olson M.V., Eichler E.E., Green E.D.,
RA   Waterston R.H., Wilson R.K.;
RT   "The DNA sequence of human chromosome 7.";
RL   Nature 424:157-164(2003).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Armstrong J., Doebber A., Haakenson W., Meyer R., VanBrunt A.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Waterston R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Wilson R.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC092180; AAS07430.1; -.
KW   Hypothetical protein.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1407 MW;  8234E644327988F7 CRC64;

  Query Match             28.2%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 1.1e+04;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 VVDKNTG 11
              |||  ||
Db          9 VVDCGTG 15


RESULT 13
Q6SJX6
ID   Q6SJX6      PRELIMINARY;      PRT;    12 AA.
AC   Q6SJX6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Acyl-protein synthetase (Fragment).
GN   Name=luxE;
OS   Photobacterium mondopomensis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Photobacterium.
OX   NCBI_TaxID=48408;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Seaf1.1.1;
RX   PubMed=15034641;
RA   Ast J.C., Dunlap P.V.;
RT   "Phylogenetic analysis of the lux operon distinguishes two
RT   evolutionarily distinct clades of Photobacterium leiognathi.";
RL   Arch. Microbiol. 181:352-361(2004).
DR   EMBL; AY456751; AAS16492.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1365 MW;  5266E00A93A05AA7 CRC64;

  Query Match             26.9%;  Score 21;  DB 2;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+04;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 VVDKNTGDI 13
              ::| :| ||
Db          4 LLDIDTNDI 12
```

```
RESULT 14
Q9UPE5
ID   Q9UPE5      PRELIMINARY;      PRT;    13 AA.
AC   Q9UPE5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Inosine monophosphatase 2 (Fragment).
GN   Name=IMPA2;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97463449; PubMed=9322233;
RA   Yoshikawa T., Turner G., Esterling L.E., Sanders A.R.,
RA   Detera-Wadleigh S.D.;
RT   "A novel human myo-inositol monophosphatase gene, IMP.18p, maps to a
RT   susceptibility region for bipolar disorder.";
RL   Mol. Psychiatry 2:393-397(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20284187;
RA   Yoshikawa T., Padigaru M., Karkera J.D., Sharma M., Berrettini W.H.,
RA   Esterling L.E., Detera-Wadleigh S.D.;
RT   "Genomic structure and novel variants of myo-inositol monophosphatase
RT   2.";
RL   Mol. Psychiatry 5:165-171(2000).
DR   EMBL; AF025886; AAD22138.1; -.
DR   EMBL; AF025885; AAD22138.1; JOINED.
FT   NON_TER       1      1
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1331 MW;  89C724C8E3457865 CRC64;

  Query Match           26.9%;  Score 21;  DB 2;  Length 13;
  Best Local Similarity   27.3%;  Pred. No. 1.3e+04;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 VVDKNTGDINI 15
              |:| : | :::
Db          2 VIDTSGGPLDL 12


RESULT 15
Q6LBZ3
ID   Q6LBZ3      PRELIMINARY;      PRT;    13 AA.
AC   Q6LBZ3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Luciferase (Fragment).
GN   Name=luxB;
OS   Vibrio harveyi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=669;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B392;
RX   MEDLINE=83117801; PubMed=6571986;
RA   Cohn D.H., Ogden R.C., Abelson J.N., Baldwin T.O., Nealson K.H.,
RA   Simon M.I., Mileham A.J.;
RT   "Cloning of the Vibrio harveyi luciferase genes: use of a synthetic
RT   oligonucleotide probe.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:120-123(1983).
DR   EMBL; V01424; CAA24694.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1596 MW;  C59F8B1F3C7E29C9 CRC64;

  Query Match           26.9%;  Score 21;  DB 2;  Length 13;
  Best Local Similarity   42.9%;  Pred. No. 1.3e+04;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FGIFVVD 7
              ||:| ::
Db          3 FGLFFLN 9


Search completed: February 22, 2005, 09:37:48
Job time : 55.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-2
Perfect score: 78
Sequence: 1 FGIFVVDKNTGDINI 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	78	100.0	15	2	AAW04842	Aaw04842 Self epit
2	46	59.0	11	3	AAY61458	Aay61458 Cadherin-
3	42	53.8	10	3	AAY61455	Aay61455 Cadherin-
4	42	53.8	11	3	AAY61477	Aay61477 Cadherin-
5	42	53.8	11	3	AAY61636	Aay61636 Cadherin-
6	40	51.3	9	3	AAY61486	Aay61486 Cadherin-
7	40	51.3	9	3	AAY61370	Aay61370 Cadherin-
8	40	51.3	9	3	AAY61678	Aay61678 Cadherin-
9	40	51.3	9	5	ABB45631	Abb45631 Non-class
10	40	51.3	11	3	AAY61438	Aay61438 Cadherin-
11	40	51.3	11	3	AAY61413	Aay61413 Cadherin-
12	40	51.3	11	3	AAY61428	Aay61428 Cadherin-
13	39	50.0	11	3	AAY63733	Aay63733 Desmoglei
14	39	50.0	11	3	AAY61076	Aay61076 Cadherin-
15	39	50.0	11	5	ABB45785	Abb45785 Desmoglei
16	38	48.7	10	3	AAY63730	Aay63730 Desmoglei
17	38	48.7	10	3	AAY61474	Aay61474 Cadherin-
18	38	48.7	10	3	AAY61633	Aay61633 Cadherin-
19	38	48.7	10	5	ABB45782	Abb45782 Desmoglei
20	38	48.7	11	3	AAY62189	Aay62189 Cadherin-
21	37	47.4	9	8	ADQ13125	Adq13125 Hepatitis
22	37	47.4	9	8	ADQ13126	Adq13126 Hepatitis
23	37	47.4	10	3	AAY61073	Aay61073 Cadherin-
24	37	47.4	10	5	ABJ08003	Abj08003 Hepatitis
25	37	47.4	10	5	ABJ08418	Abj08418 Hepatitis
26	37	47.4	10	8	ADK38959	Adk38959 Hepatitis
27	37	47.4	11	5	ABJ08409	Abj08409 Hepatitis
28	37	47.4	11	5	ABJ07987	Abj07987 Hepatitis
29	37	47.4	15	5	ABJ09323	Abj09323 Hepatitis
30	37	47.4	15	5	ABJ08966	Abj08966 Hepatitis
31	37	47.4	15	5	ABJ09353	Abj09353 Hepatitis
32	37	47.4	15	5	ABJ09974	Abj09974 Hepatitis
33	37	47.4	15	5	ABJ09258	Abj09258 Hepatitis
34	37	47.4	15	8	ADK39235	Adk39235 Hepatitis
35	37	47.4	15	8	ADK39349	Adk39349 Hepatitis
36	36	46.2	8	3	AAY61675	Aay61675 Cadherin-
37	36	46.2	8	3	AAY61367	Aay61367 Cadherin-
38	36	46.2	8	3	AAY61483	Aay61483 Cadherin-
39	36	46.2	9	3	AAY61452	Aay61452 Cadherin-
40	36	46.2	10	3	AAY61410	Aay61410 Cadherin-
41	36	46.2	10	3	AAY61427	Aay61427 Cadherin-
42	36	46.2	10	3	AAY61437	Aay61437 Cadherin-
43	36	46.2	11	3	AAY61965	Aay61965 Cadherin-
44	36	46.2	11	3	AAY63718	Aay63718 Desmoglei
45	36	46.2	11	5	ABB45770	Abb45770 Desmoglei

ALIGNMENTS

RESULT 1
AAW04842
ID AAW04842 standard; peptide; 15 AA.
XX
AC AAW04842;
XX
DT 18-FEB-1997 (first entry)
XX
DE Self epitope of desmoglein 3, implicated in autoimmune disease.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Homo sapiens.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 1; Page 38; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is derived from the human desmoglein

```
CC   3 protein (amino acids 97-111) and is implicated as a self epitope in
CC   pemphigus vulgaris. Peptides derived from the human desmoglein protein
CC   are described in AAW04841-47
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 78;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 FGIFVVDKNTGDINI 15
              |||||||||||||||
Db          1 FGIFVVDKNTGDINI 15


RESULT 2
AAY61458
ID   AAY61458 standard; peptide; 11 AA.
XX
AC   AAY61458;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1343.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 172; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             59.0%;  Score 46;  DB 3;  Length 11;
  Best Local Similarity   70.0%;  Pred. No. 0.34;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTGD 12
              ||::|:||||
Db          2 IFIIDENTGD 11


RESULT 3
AAY61455
ID   AAY61455 standard; peptide; 10 AA.
XX
AC   AAY61455;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1340.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .10
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
```

```
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 172; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 10 AA;

  Query Match             53.8%;  Score 42;  DB 3;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 1.6;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 FVVDKNTGD 12
              |::|:||||
Db          2 FIIDENTGD 10


RESULT 4
AAY61477
ID   AAY61477 standard; peptide; 11 AA.
XX
AC   AAY61477;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1363.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 172; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             53.8%;  Score 42;  DB 3;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 1.7;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTGD 12
              ||::|:|||:
Db          2 IFIIDENTGE 11


RESULT 5
AAY61636
ID   AAY61636 standard; peptide; 11 AA.
XX
AC   AAY61636;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1522.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;    99WO-CA000363.
XX
PR   05-MAY-1998;    98US-00073040.
PR   06-NOV-1998;    98US-00187859.
PR   20-JAN-1999;    99US-00234395.
PR   08-MAR-1999;    99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 174; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             53.8%;  Score 42;  DB 3;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 1.7;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTGD 12
              ||::|:|||:
Db          2 IFIIDENTGE 11


RESULT 6
AAY61486
ID   AAY61486 standard; peptide; 9 AA.
XX
AC   AAY61486;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1372.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .9
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;    99WO-CA000363.
XX
PR   05-MAY-1998;    98US-00073040.
PR   06-NOV-1998;    98US-00187859.
PR   20-JAN-1999;    99US-00234395.
PR   08-MAR-1999;    99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 172; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 9 AA;

  Query Match             51.3%;  Score 40;  DB 3;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTG 11
              ||::|:|||
Db          1 IFIIDENTG 9
```

```
RESULT 7
AAY61370
ID   AAY61370 standard; peptide; 9 AA.
XX
AC   AAY61370;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition sequence SEQ ID NO:196.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 33; Page 170; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 9 AA;

  Query Match             51.3%;  Score 40;  DB 3;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTG 11
              ||::|:|||
Db          1 IFIIDENTG 9


RESULT 8
AAY61678
ID   AAY61678 standard; peptide; 9 AA.
XX
AC   AAY61678;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1564.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .9
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 175; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
```

```
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 9 AA;

  Query Match             51.3%;  Score 40;  DB 3;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTG 11
              ||::|:|||
Db          1 IPIIDENTG 9


RESULT 9
ABB45631
ID   ABB45631 standard; peptide; 9 AA.
XX
AC   ABB45631;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Non-classical cadherin CAR sequence SEQ ID NO 375.
XX
KW   Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;
KW   cytostatic; antiapoptotic; wound healing; reduce scar tissue; skin graft;
KW   organ implant; autoimmune blistering disorder; cancer; apoptosis.
XX
OS   Synthetic.
XX
PN   WO200172956-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-IB001400.
XX
PR   27-MAR-2000; 2000US-00535852.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Symonds JM,  Gour BJ;
XX
DR   WPI; 2002-025778/03.
XX
PT   Modulating agents for inhibiting or enhancing desmosomal cadherin
PT   mediated cell adhesion, useful for facilitating wound healing and/or
PT   reducing scar tissue, treating cancer and inducing apoptosis.
XX
PS   Disclosure; Page 21; 127pp; English.
XX
CC   The invention relates to modulating agents for inhibiting or enhancing
CC   desmosomal cadherin mediated cell adhesion, comprising a modulating agent
CC   comprising a desmosomal cadherin cell adhesion recognition CAR sequence
CC   (ABB45341-ABB47262), a non-peptide mimetic of a desmosomal cadherin CAR
CC   sequence, a substance such as an antibody or antigen-binding fragment
CC   that specifically binds a desmosomal cadherin CAR sequence and/or a
CC   polynucleotide encoding a polypeptide that comprises a desmosomal
CC   cadherin CAR sequence or analogue. The modulating agents have
CC   immunosuppressive, cytostatic and antiapoptotic activity and are used to
CC   facilitate wound healing and/or reduce scar tissue, for enhancing
CC   adhesion of foreign tissue implants (e.g. skin graft or organ implant),
CC   treating an autoimmune blistering disorder and to treat cancer (e.g.
CC   carcinoma, leukaemia or melanoma) and induce apoptosis
XX
SQ   Sequence 9 AA;

  Query Match             51.3%;  Score 40;  DB 5;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTG 11
              ||::|:|||
Db          1 IPIIDENTG 9


RESULT 10
AAY61438
ID   AAY61438 standard; peptide; 11 AA.
XX
AC   AAY61438;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:3913.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 171; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
```

```
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             51.3%;  Score 40;  DB 3;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 3.9;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTG 11
              ||::|:|||
Db          2 IFIIDENTG 10


RESULT 11
AAY61413
ID   AAY61413 standard; peptide; 11 AA.
XX
AC   AAY61413;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:1317.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  1. .11
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 171; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             51.3%;  Score 40;  DB 3;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 3.9;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTG 11
              ||::|:|||
Db          2 IFIIDENTG 10


RESULT 12
AAY61428
ID   AAY61428 standard; peptide; 11 AA.
XX
AC   AAY61428;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-7 cell adhesion recognition cyclic peptide SEQ ID NO:3903.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
```

```
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 36; Page 171; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             51.3%;  Score 40;  DB 3;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 3.9;
  Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTG 11
              ||::|:|||
Db          2 IFIIDENTG 10


RESULT 13
AAY63733
ID   AAY63733 standard; peptide; 11 AA.
XX
AC   AAY63733;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Desmoglein cell adhesion recognition cyclic peptide SEQ ID NO:3185.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 90; Page 208; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             50.0%;  Score 39;  DB 3;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 5.8;
  Matches    5;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTGD 12
              :|::::||||
Db          2 MFIINRNTGD 11


RESULT 14
AAY61076
ID   AAY61076 standard; peptide; 11 AA.
XX
AC   AAY61076;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Cadherin-6 cell adhesion recognition cyclic peptide SEQ ID NO:1022.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
```

```
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 27; Page 166; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             50.0%;  Score 39;  DB 3;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 5.8;
  Matches    5;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTGD 12
              :|::::||||
Db          2 LFIINENTGD 11


RESULT 15
ABB45785
ID   ABB45785 standard; peptide; 11 AA.
XX
AC   ABB45785;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Desmoglein CAR sequence cyclic peptide SEQ ID NO 529.
XX
KW   Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;
KW   cytostatic; antiapoptotic; wound healing; reduce scar tissue; skin graft;
KW   organ implant; autoimmune blistering disorder; cancer; apoptosis; cyclic.
XX
OS   Synthetic.
XX
PN   WO200172956-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-IB001400.
XX
PR   27-MAR-2000; 2000US-00535852.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Symonds JM,  Gour BJ;
XX
DR   WPI; 2002-025778/03.
XX
PT   Modulating agents for inhibiting or enhancing desmosomal cadherin
PT   mediated cell adhesion, useful for facilitating wound healing and/or
PT   reducing scar tissue, treating cancer and inducing apoptosis.
XX
PS   Claim 18; Page 98; 127pp; English.
XX
CC   The invention relates to modulating agents for inhibiting or enhancing
CC   desmosomal cadherin mediated cell adhesion, comprising a modulating agent
CC   comprising a desmosomal cadherin cell adhesion recognition CAR sequence
CC   (ABB45341-ABB47262), a non-peptide mimetic of a desmosomal cadherin CAR
CC   sequence, a substance such as an antibody or antigen-binding fragment
CC   that specifically binds a desmosomal cadherin CAR sequence and/or a
CC   polynucleotide encoding a polypeptide that comprises a desmosomal
CC   cadherin CAR sequence or analogue. The modulating agents have
CC   immunosuppressive, cytostatic and antiapoptotic activity and are used to
CC   facilitate wound healing and/or reduce scar tissue, for enhancing
CC   adhesion of foreign tissue implants (e.g. skin graft or organ implant),
CC   treating an autoimmune blistering disorder and to treat cancer (e.g.
CC   carcinoma, leukaemia or melanoma) and induce apoptosis
XX
SQ   Sequence 11 AA;

  Query Match             50.0%;  Score 39;  DB 5;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 5.8;
  Matches    5;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 IFVVDKNTGD 12
              :|::::||||
Db          2 MFIINRNTGD 11


Search completed: February 22, 2005, 09:24:33
Job time : 66.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-3
Perfect score: 71
Sequence: 1 LNSKIAFKIVSQEPA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	35.2	10	2	S66458	ferredoxin - Rhizo
2	23	32.4	14	2	A49018	myosin heavy chain
3	20	28.2	14	2	PH0795	T-cell receptor al
4	20	28.2	14	2	PH0776	T-cell receptor al
5	20	28.2	15	2	PA0097	starch phosphoryla
6	19	26.8	10	2	F44644	neurotoxin-associa
7	19	26.8	12	2	S71034	potB protein - Sal
8	19	26.8	13	2	S57571	T cell receptor al
9	19	26.8	13	2	E42762	proteasome endopep
10	19	26.8	13	2	G44644	neurotoxin-associa
11	19	26.8	15	2	PS0452	32K protein 3306 -
12	18	25.4	10	2	PT0309	Ig heavy chain CRD
13	18	25.4	10	2	A24407	amicyanin - Paraco
14	18	25.4	11	2	E60691	phycobilisome 8K l
15	18	25.4	12	2	D28551	hypothetical prote
16	18	25.4	12	2	S27024	Na+/K+-exchanging
17	18	25.4	14	2	PH0801	T-cell receptor al
18	18	25.4	15	2	PQ0174	stylar glycoprotei
19	18	25.4	15	2	PS0251	15K protein 5106 -
20	18	25.4	15	2	PA0054	protein QF200017 -
21	18	25.4	15	2	A32971	heparin-binding le
22	17.5	24.6	15	2	PT0222	Ig heavy chain CDR
23	17	23.9	3	3	A33802	thyrotropin-releas
24	17	23.9	10	2	PH0807	T-cell receptor al
25	17	23.9	11	2	G61497	seed protein ws-23
26	17	23.9	12	2	A55837	5-aminoimidazole r
27	17	23.9	13	2	H64124	hypothetical prote
28	17	23.9	13	2	PL0157	Ig kappa chain V-I
29	17	23.9	13	2	B56864	dipeptidyl-peptida
30	17	23.9	13	2	A59387	VCAM-1 5'UTR bindi
31	17	23.9	14	2	S35267	acetyl-CoA carboxy
32	17	23.9	14	2	JH0516	insulin-like growt
33	17	23.9	15	2	PA0046	protein QA100044 -
34	17	23.9	15	2	PH0782	T-cell receptor al
35	17	23.9	15	2	S77987	cytochrome-c oxida
36	16	22.5	8	2	S37141	rpsA protein - Erw
37	16	22.5	9	2	D48186	ATPase R1 subunit
38	16	22.5	9	2	I54379	gene NF2 protein -
39	16	22.5	9	2	PC7073	ubiquinol-cytochro
40	16	22.5	10	2	B39308	glycine reductase
41	16	22.5	11	1	EOOC	eledoisin - musky
42	16	22.5	11	1	EOOCC	eledoisin - curled
43	16	22.5	11	2	S53436	beta-D-galactosida
44	16	22.5	11	2	PQ0231	beta-glucosidase (
45	16	22.5	12	2	A29169	phospholipase A2 (

ALIGNMENTS

RESULT 1
S66458
ferredoxin - Rhizobium meliloti (fragment)
C;Species: Rhizobium meliloti
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S66458
R;Riedel, K.U.; Jouanneau, Y.; Masepohl, B.; Puehler, A.; Klipp, W.
Eur. J. Biochem. 231, 742-746, 1995
A;Title: A Rhizobium meliloti ferredoxin (FdxN) purified from Escherichia coli donates
A;Reference number: S66458; MUID:95377307; PMID:7649175
A;Accession: S66458
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <RIE>
C;Genetics:
A;Gene: fdxN

Query Match 35.2%; Score 25; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          6 AFKIVSQE 13
              |||:: :
Db          1 AFKIIASQ 8
```

RESULT 2
A49018
myosin heavy chain, fast skeletal muscle type X - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 13-Aug-1999
C;Accession: A49018; S32161
R;DeNardi, C.; Ausoni, S.; Moretti, P.; Gorza, L.; Velleca, M.; Buckingham, M.; Schiaffi
J. Cell Biol. 123, 823-835, 1993
A;Title: Type 2X-myosin heavy chain is coded by a muscle fiber type-specific and develo
A;Reference number: A49018; MUID:94043465; PMID:8227143
A;Accession: A49018
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-14 <DEN>
A;Cross-references: GB:X72591; NID:g288645; PIDN:CAA51189.1; PID:g288646
R;DeNardi, C.; Ausoni, S.; Moretti, P.; Gorza, L.; Velleca, M.; Merlie, J.; Buckingham,
submitted to the EMBL Data Library, March 1993
A;Description: Type 2X myosin heavy chain is coded by a muscle fiber type-specific and
A;Reference number: S32161
A;Accession: S32161
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-14 <DEN2>
A;Cross-references: EMBL:X72591; NID:g288645; PIDN:CAA51189.1; PID:g288646
C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: skeletal muscle

```
  Query Match             32.4%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 6.7e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 KIVSQE 13
              ||:|:|
Db          9 KIISEE 14


RESULT 3
PH0795
T-cell receptor alpha chain (K1 V-alpha-4.3) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0795
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0795
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60900
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             28.2%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 NSKIAF 7
              |:|: |
Db          8 NAKLTF 13


RESULT 4
PH0776
T-cell receptor alpha chain (M1 V-alpha-8.F3.3) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0776
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0776
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60873
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             28.2%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 NSKIAF 7
              |:|: |
Db          8 NNKLTF 13


RESULT 5
PA0097
starch phosphorylase (EC 2.4.1.1) - fungus (Fusarium sporotrichioides) (fragment)
N;Alternate names: maltodextrin phosphorylase
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C;Accession: PA0097
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0097
A;Molecule type: protein
A;Residues: 1-15 <CHO>
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             28.2%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 SKIAFKIVSQ 12
              :|  | |  |
Db          2 NKFVFTITGQ 11


RESULT 6
F44644
neurotoxin-associated protein type B Hn+ 35K chain, band 3a - Clostridium botulinum (fra
C;Species: Clostridium botulinum
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: F44644
R;Somers, E.; DasGupta, B.R.
J. Protein Chem. 10, 415-425, 1991
A;Title: Clostridium botulinum types A, B, C1, and E produce proteins with or without he
A;Reference number: A44644; MUID:92143938; PMID:1781887
A;Contents: type B
A;Accession: F44644
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <SOM>
A;Cross-references: UNIPROT:Q9R5N6
A;Note: sequence extracted from NCBI backbone (NCBIP:83787)
C;Keywords: hemagglutinin

  Query Match             26.8%;  Score 19;  DB 2;  Length 10;
  Best Local Similarity   80.0%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LNSKI 5
              || ||
Db          5 LNDKI 9


RESULT 7
S71034
potB protein - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S71034
R;Stein, M.A.; Leung, K.Y.; Zwick, M.; Garcia-del Portillo, F.; Finlay, B.B.
Mol. Microbiol. 20, 151-164, 1996
A;Title: Identification of a Salmonella virulence gene required for formation of filamen
A;Reference number: S71033; MUID:97014378; PMID:8861213
A;Accession: S71034
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-12 <STE>
A;Cross-references: UNIPROT:Q56060; EMBL:U51867; NID:g1272352; PIDN:AAA97466.1; PID:g127
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1996
C;Genetics:
A;Gene: potB

  Query Match             26.8%;  Score 19;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 3.2e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LNSKIA 6
              || |::
Db          3 LNKKVS 8


RESULT 8
```

```
S57571
T cell receptor alpha chain V-J region (clone PP7 and others) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57571; S57573; S57576
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A;Reference number: S57494
A;Accession: S57571
A;Molecule type: mRNA
A;Residues: 1-13 <BUR>
A;Cross-references: EMBL:Z49948; NID:g887496; PIDN:CAA90219.1; PID:g887497
A;Experimental source: clone PP7
A;Accession: S57573
A;Molecule type: mRNA
A;Residues: 1-13 <BUL>
A;Cross-references: EMBL:Z49950; NID:g887500; PIDN:CAA90221.1; PID:g887501
A;Experimental source: clone TF1
A;Accession: S57576
A;Molecule type: mRNA
A;Residues: 1-13 <BUW>
A;Cross-references: EMBL:Z49952; NID:g887512; PIDN:CAA90223.1; PID:g887513
A;Experimental source: clone RL16
C;Keywords: T-cell receptor

  Query Match             26.8%;  Score 19;  DB 2;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 3.5e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NSKIAF 7
              | |: |
Db          7 NEKLTF 12


RESULT 9
E42762
proteasome endopeptidase complex (EC 3.4.25.1) subunit 13 - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 17-Feb-2003
C;Accession: E42762
R;Dick, L.R.; Moomaw, C.R.; Pramanik, B.C.; DeMartino, G.N.; Slaughter, C.A.
Biochemistry 31, 7347-7355, 1992
A;Title: Identification and localization of a cysteinyl residue critical for the trypsin
A;Reference number: A42762; MUID:92378961; PMID:1510924
A;Accession: E42762
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-13 <DIC>
A;Note: sequence extracted from NCBI backbone (NCBIP:112180)
C;Keywords: hydrolase

  Query Match             26.8%;  Score 19;  DB 2;  Length 13;
  Best Local Similarity   25.0%;  Pred. No. 3.5e+03;
  Matches    3;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 LNSKIAFKIVSQ 12
              :: :  |: :||
Db          1 MDPEXLFETISQ 12


RESULT 10
G44644
neurotoxin-associated protein type B Hn+ 35K chain, band 3b - Clostridium botulinum (fra
C;Species: Clostridium botulinum
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: G44644
R;Somers, E.; DasGupta, B.R.
J. Protein Chem. 10, 415-425, 1991
A;Title: Clostridium botulinum types A, B, C1, and E produce proteins with or without he
A;Reference number: A44644; MUID:92143938; PMID:1781887
A;Contents: type B
A;Accession: G44644
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-13 <SOM>
A;Cross-references: UNIPROT:Q9R5N5
A;Note: sequence extracted from NCBI backbone (NCBIP:83785)
C;Keywords: hemagglutinin

  Query Match             26.8%;  Score 19;  DB 2;  Length 13;
  Best Local Similarity   80.0%;  Pred. No. 3.5e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LNSKI 5
              || ||
Db          8 LNDKI 12


RESULT 11
PS0452
32K protein 3306 - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 11-Apr-1995
C;Accession: PS0452
R;Tsugita, A.; Miyatake, N.
submitted to JIPID, April 1993
A;Reference number: PS0208
A;Accession: PS0452
A;Molecule type: protein
A;Residues: 1-15 <TSU>
A;Experimental source: bran, strain Nihonbare
C;Comment: molecular weight 32K, pI 5.3.

  Query Match             26.8%;  Score 19;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 4.1e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 FKIVSQEP 14
              | |: | |
Db          7 FPILIQXP 14


RESULT 12
PT0309
Ig heavy chain CRD3 region (clone 6-94) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0309
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0309
A;Molecule type: DNA
A;Residues: 1-10 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

  Query Match             25.4%;  Score 18;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 NSKIAF 7
              ||: |:
Db          5 NSRAAY 10


RESULT 13
A24407
amicyanin - Paracoccus denitrificans (fragment)
C;Species: Paracoccus denitrificans
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 09-Jul-2004
C;Accession: A24407
R;Husain, M.; Davidson, V.L.
Biochemistry 25, 2431-2436, 1986
```

```
A;Title: Properties of Paracoccus denitrificans amicyanin.
A;Reference number: A24407; MUID:86243362; PMID:3718960
A;Accession: A24407
A;Molecule type: protein
A;Residues: 1-10 <HUS>
A;Cross-references: UNIPROT:P22364

  Query Match             25.4%;  Score 18;  DB 2;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 4.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 IVSQEP 14
              | || |
Db          5 IPSQSP 10


RESULT 14
E60691
phycobilisome 8K linker protein - Synechococcus sp. (PCC 7002) (fragment)
C;Species: Synechococcus sp.
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 07-May-1999
C;Accession: E60691
R;Bryant, D.A.; de Lorimier, R.; Guglielmi, G.; Stevens Jr., S.E.
Arch. Microbiol. 153, 550-560, 1990
A;Title: Structural and compositional analyses of the phycobilisomes of Synechococcus sp
s.
A;Reference number: A60691; MUID:90314662; PMID:2164365
A;Accession: E60691
A;Molecule type: protein
A;Residues: 1-11 <BRY>
C;Comment: This protein, one of the eleven components detected in this species of the ph
C;Keywords: photosystem II

  Query Match             25.4%;  Score 18;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 4.5e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 FKIVSQEP 14
              |||  :  |
Db          4 FKITACVP 11


RESULT 15
D28551
hypothetical protein 4 - Streptococcus mutans (strain GS-5) (fragment)
C;Species: Streptococcus mutans
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 15-Oct-1999
C;Accession: D28551
R;Shiroza, T.; Kuramitsu, H.K.
J. Bacteriol. 170, 810-816, 1988
A;Title: Sequence analysis of the Streptococcus mutans fructosyltransferase gene and fla
A;Reference number: A91892; MUID:88115184; PMID:2828325
A;Accession: D28551
A;Molecule type: DNA
A;Residues: 1-12 <SHI>
A;Cross-references: GB:M18954; NID:g153635; PIDN:AAA88586.1; PID:g1196927

  Query Match             25.4%;  Score 18;  DB 2;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 5e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 NSKIAFKI 9
              | ||  ::
Db          3 NKKIRIRL 10


Search completed: February 22, 2005, 09:46:24
Job time : 12.1333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-3
Perfect score: 71
Sequence: 1 LNSKIAFKIVSQEPA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	36.6	15	2	Q6GWV7	Q6gwv7 bos taurus
2	24	33.8	13	1	PROX_ORYSA	P83647 oryza sativ
3	23	32.4	13	2	Q957T7	Q957t7 trichopsis
4	23	32.4	14	2	Q06415	Q06415 rattus ratt
5	22	31.0	9	2	Q70YA2	Q70ya2 alvesia ros
6	22	31.0	14	1	UC15_MAIZE	P80621 zea mays (m
7	22	31.0	14	2	Q06414	Q06414 rattus ratt
8	22	31.0	15	1	THBI_BOTJA	P84026 bothrops ja
9	21	29.6	15	2	Q91XV8	Q91xv8 rattus norv
10	20	28.2	9	2	Q45533	Q45533 bacillus su
11	20	28.2	10	2	Q9QVK8	Q9qvk8 mus sp. mep
12	20	28.2	11	2	Q34380	Q34380 drosophila
13	20	28.2	13	2	Q9ZEZ1	Q9zez1 buchnera ap
14	20	28.2	14	2	Q16118	Q16118 homo sapien
15	20	28.2	14	2	Q6SE52	Q6se52 drosophila
16	20	28.2	14	2	O93202	O93202 porcine cir
17	20	28.2	14	2	Q77NR3	Q77nr3 porcine cir
18	20	28.2	14	2	Q77RC0	Q77rc0 porcine cir
19	20	28.2	14	2	Q77RM6	Q77rm6 porcine cir
20	20	28.2	14	2	Q77S02	Q77s02 porcine cir
21	20	28.2	14	2	Q77S09	Q77s09 porcine cir
22	20	28.2	14	2	Q77S13	Q77s13 bovine circ
23	19	26.8	10	1	CX81_CANFA	P61904 canis famil
24	19	26.8	10	2	Q83YL7	Q83yl7 acinetobact
25	19	26.8	10	2	Q9R5N4	Q9r5n4 clostridium
26	19	26.8	10	2	Q9R5N6	Q9r5n6 clostridium
27	19	26.8	11	2	Q9L4F7	Q914f7 bacillus ce
28	19	26.8	13	1	BP37_LEUMA	P81754 leucophaea
29	19	26.8	13	2	Q93A39	Q93a39 enterococcu
30	19	26.8	13	2	Q55234	Q55234 synechocyst
31	19	26.8	13	2	Q9R5N5	Q9r5n5 clostridium
32	19	26.8	14	2	Q9BRY8	Q9bry8 homo sapien
33	19	26.8	14	2	Q7RP61	Q7rp61 plasmodium
34	19	26.8	14	2	Q6GWV6	Q6gwv6 ovis aries
35	19	26.8	14	2	Q9FUX5	Q9fux5 symphoricar
36	19	26.8	15	1	MP2A_ORYSA	P83466 oryza sativ
37	19	26.8	15	2	P82211	P82211 bombyx mori
38	18	25.4	10	2	Q9R7J8	Q9r7j8 helicobacte
39	18	25.4	10	2	Q9JLI5	Q9jli5 mus musculu
40	18	25.4	11	2	Q56413	Q56413 escherichia
41	18	25.4	12	2	Q79DY3	Q79dy3 streptococc
42	18	25.4	13	2	P82276	P82276 homo sapien
43	18	25.4	13	2	O82835	O82835 synechococc
44	18	25.4	13	2	Q918T4	Q918t4 human papil
45	18	25.4	13	2	Q918T6	Q918t6 human papil

ALIGNMENTS

RESULT 1
Q6GWV7
ID Q6GWV7 PRELIMINARY; PRT; 15 AA.
AC Q6GWV7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Interleukin 8 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Meade K.G., Fitzgerald D.C., Murphy E.P., Baird A.W., MacHugh D.E.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY627308; AAT47550.1; -.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1504 MW; DA1932E93C487C14 CRC64;

Query Match 36.6%; Score 26; DB 2; Length 15;
Best Local Similarity 33.3%; Pred. No. 1.2e+03;
Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 LNSKIAFKIVSQEPA 15
              : ||:|  :::  ||
Db          1 MTSKLAVALLAAFPA 15
```

RESULT 2
PROX_ORYSA
ID PROX_ORYSA STANDARD; PRT; 13 AA.
AC P83647;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Probable profilin LP04 (Fragments).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Indica / IR64; TISSUE=Panicle;
RA Hosseini Salekdeh S.G., Bennett J.;
RT "Proteome analysis of rice panicle.";
RL Submitted (JUL-2003) to Swiss-Prot.
CC -!- FUNCTION: Binds to actin and affects the structure of the
CC cytoskeleton. At high concentrations, profilin prevents the
CC polymerization of actin, whereas it enhances it at low
CC concentrations. By binding to PIP2, it inhibits the formation of
CC IP3 and DG (By similarity).

```
CC   -!- SUBUNIT: Occurs in many kinds of cells as a complex with monomeric
CC       actin in a 1:1 ratio.
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 4.4, its MW is: 14.0 kDa.
CC   -!- SIMILARITY: Belongs to the profilin family.
DR   InterPro; IPR002097; Profilin.
DR   PROSITE; PS00414; PROFILIN; PARTIAL.
KW   Actin-binding; Cytoskeleton; Direct protein sequencing;
KW   Multigene family.
FT   NON_TER       1      1
FT   NON_CONS      5      6
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1362 MW;  0A3022EE0E52C68B CRC64;

  Query Match             33.8%;  Score 24;  DB 1;  Length 13;
  Best Local Similarity   30.0%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 IAFKIVSQEP 14
              :|: :: ||
Db          4 LAYMVIQGEP 13


RESULT 3
Q957T7
ID   Q957T7      PRELIMINARY;      PRT;    13 AA.
AC   Q957T7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome b (Fragment).
OS   Trichopsis pumila (pygmy gourami).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes;
OC   Anabantoidei; Belontiidae; Trichopsis.
OX   NCBI_TaxID=158454;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang T.Y., Tzeng C.S., Shen S.C.;
RT   "Conservation and Phylogeography of Taiwan Paradise Fish, Macropodus
RT   opercularis Linnaeus.";
RL   Acta Zool. Taiwanica 10:121-134(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Wang T.Y., Tzeng C.S., Shen S.C.;
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF359378; AAK51458.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1443 MW;  55430C67982AAB17 CRC64;

  Query Match             32.4%;  Score 23;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 3.8e+03;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 LNSKIAFK 8
              |  || ||
Db          6 LEDKILFK 13


RESULT 4
Q06415
ID   Q06415      PRELIMINARY;      PRT;    14 AA.
AC   Q06415;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   2X myosin heavy chain (Fragment).
OS   Rattus rattus (Black rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10117;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Diaphragm;
RX   MEDLINE=94043465; PubMed=8227143; DOI=10.1083/jcb.123.4.823;
RA   DeNardi C., Ausoni S., Moretti P., Gorza L., Velleca M.,
RA   Buckingham M., Schiaffino S.;
RT   "Type 2X Myosin Heavy Chain is coded by a muscle fiber type-specific
RT   and developmentally regulated gene.";
RL   J. Cell Biol. 123:823-835(1993).
DR   EMBL; X72591; CAA51189.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1655 MW;  F1C0536CD7B6AFA1 CRC64;

  Query Match             32.4%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 4.1e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 KIVSQE 13
              ||:|:|
Db          9 KIISEE 14


RESULT 5
Q70YA2
ID   Q70YA2      PRELIMINARY;      PRT;     9 AA.
AC   Q70YA2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Ribosomal protein (Fragment).
GN   Name=rps16;
OS   Alvesia rosmarinifolia.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Alvesia.
OX   NCBI_TaxID=204103;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15019625; DOI=10.1016/j.ympev.2003.08.002;
RA   Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Harley M.M.,
RA   Willis F., Simonds M.S.J., Powell M.P., Savolainen V.;
RT   "Phylogeny and evolution of basils and allies (Ocimeae, Labiatae)
RT   based on three plastid DNA regions.";
RL   Mol. Phylogenet. Evol. 31:277-299(2004).
DR   EMBL; AJ505329; CAD45452.1; -.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Ribosomal protein.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  979 MW;  0DF006C5B721F2C0 CRC64;

  Query Match             31.0%;  Score 22;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 VSQEP 14
              :||||
Db          5 LSQEP 9


RESULT 6
UC15_MAIZE
ID   UC15_MAIZE     STANDARD;      PRT;    14 AA.
AC   P80621;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Unknown protein from 2D-PAGE of etiolated coleoptile (Spot 245)
DE   (Fragment).
OS   Zea mays (Maize).
```

```
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Coleoptile;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 4.8, its MW is: 35.7 kDa.
DR   Maize-2DPAGE; P80621; COLEOPTILE.
DR   MaizeDB; 123947; -.
KW   Direct protein sequencing.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1396 MW;  C68949275F404CD2 CRC64;

  Query Match             31.0%;  Score 22;  DB 1;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 6.4e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 IVSQEPA 15
              : ::|||
Db          8 VAAEEPA 14


RESULT 7
Q06414
ID   Q06414      PRELIMINARY;      PRT;    14 AA.
AC   Q06414; Q06918;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   2B myosin heavy chain (2A myosin heavy chain) (Fragment).
OS   Rattus rattus (Black rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10117;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Diaphragm;
RX   MEDLINE=94043465; PubMed=8227143; DOI=10.1083/jcb.123.4.823;
RA   DeNardi C., Ausoni S., Moretti P., Gorza L., Velleca M.,
RA   Buckingham M., Schiaffino S.;
RT   "Type 2X Myosin Heavy Chain is coded by a muscle fiber type-specific
RT   and developmentally regulated gene.";
RL   J. Cell Biol. 123:823-835(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Diaphragm;
RA   Schiaffino S.;
RL   Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X72590; CAA51188.1; -.
DR   EMBL; X72589; CAA51187.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1641 MW;  F1C0536CC526AFA1 CRC64;

  Query Match             31.0%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 6.4e+03;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 KIVSQE 13
              |::|:|
Db          9 KVISEE 14


RESULT 8
THBI_BOTJA
ID   THBI_BOTJA     STANDARD;      PRT;    15 AA.
AC   P84026;
DT   25-OCT-2004 (Rel. 45, Created)
DT   25-OCT-2004 (Rel. 45, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Thrombin inhibitor subunit I (Fragment).
OS   Bothrops jararaca (Jararaca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Bothrops.
OX   NCBI_TaxID=8724;
RN   [1]
RP   SEQUENCE, FUNCTION, SUBCELLULAR LOCATION, TISSUE SPECIFICITY, SUBUNIT,
RP   AND GLYCOSYLATION.
RC   TISSUE=Plasma;
RX   PubMed=12927776; DOI=10.1016/S0006-291X(03)01464-5;
RA   Tanaka-Azevedo A.M., Tanaka A.S., Sano-Martins I.S.;
RT   "A new blood coagulation inhibitor from the snake Bothrops jararaca
RT   plasma: isolation and characterization.";
RL   Biochem. Biophys. Res. Commun. 308:706-712(2003).
CC   -!- FUNCTION: Binds to thrombin and inhibits blood coagulant activity.
CC       Does not inhibit the serine protease activity of thrombin or
CC       platelet aggregation activity. May bind to the heparin recognition
CC       exosite.
CC   -!- SUBUNIT: Heteropolymer of subunit I and subunit II.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Plasma.
CC   -!- PTM: Glycosylated.
KW   Blood coagulation; Direct protein sequencing; Glycoprotein.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1651 MW;  CE8E339D80419137 CRC64;

  Query Match             31.0%;  Score 22;  DB 1;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 6.9e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 FKIVSQEP 14
              |  |:|:|
Db          3 FPAVNQKP 10


RESULT 9
Q91XV8
ID   Q91XV8      PRELIMINARY;      PRT;    15 AA.
AC   Q91XV8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   20alpha-hydroxysteroid dehydrogenase (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Nakajima T., Yasuda K., Okada H., Sanezumi M., Osaki T., Kanzaki H.,
RA   Nishizawa, M, Ito S.;
RT   "Expression of 20alpha-HSD (hydroxysteroid dehydrogenase) in
RT   endometrium.";
RL   J. Fertilization and Implantation 17:186-187(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Nishizawa M., Nakajima T., Ito S.;
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB028066; BAB62013.1; -.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1732 MW;  B452E70F45C4822F CRC64;

  Query Match             29.6%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 LNSKI 5
              :||||
Db          1 MNSKI 5


RESULT 10
Q45533
ID   Q45533      PRELIMINARY;      PRT;     9 AA.
AC   Q45533;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=W168;
RX   MEDLINE=92041654; PubMed=1938941;
RA   Fuhrer D.K., Ordal G.W.;
RT   "Bacillus subtilis CheN, a homolog of CheA, the central regulator of
RT   chemotaxis in Escherichia coli.";
RL   J. Bacteriol. 173:7443-7448(1991).
DR   EMBL; M57894; AAA22312.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  966 MW;  87E0C6C5A761F6C5 CRC64;

  Query Match             28.2%;  Score 20;  DB 2;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+06;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 VSQEPA 15
              :|| |:
Db          1 ISQHPS 6


RESULT 11
Q9QVK8
ID   Q9QVK8      PRELIMINARY;      PRT;    10 AA.
AC   Q9QVK8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   MEPRIN=METALLOENDOPEPTIDASE (Fragment).
OS   Mus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10095;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=91363409; PubMed=1888759; DOI=10.1016/0167-4838(91)90032-U;
RA   Flannery A.V., Macadam G.C., Beynon R.J.;
RT   "Immunological characterisation of different meprin species in mice.";
RL   Biochim. Biophys. Acta 1079:119-122(1991).
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1163 MW;  DD6436144731B2C9 CRC64;

  Query Match             28.2%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   44.4%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IAFKIVSQE 13
              |||  :::|
Db          1 IAFVTLNEE 9


RESULT 12
Q34380
ID   Q34380      PRELIMINARY;      PRT;    11 AA.
AC   Q34380;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   NADH dehydrogenase subunit 1 (Fragment).
GN   Name=mt:ND1; Synonyms=ND1;
OS   Drosophila subobscura (Fruit fly).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7241;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95106285; PubMed=7528808;
RA   Barrio E., Latorre A., Moya A.;
RT   "Phylogeny of the Drosophila obscura species group deduced from
RT   mitochondrial DNA sequences.";
RL   J. Mol. Evol. 39:478-488(1994).
DR   EMBL; U07287; AAA76625.1; -.
DR   FlyBase; FBgn0012956; Dsub\mt:ND1.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   11 AA;  1264 MW;  46AEF494C7272053 CRC64;

  Query Match             28.2%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+04;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IAFKIV 10
              | |||:
Db          1 IGFKIL 6


RESULT 13
Q9ZEZ1
ID   Q9ZEZ1      PRELIMINARY;      PRT;    13 AA.
AC   Q9ZEZ1;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   2-isopropylmalate synthase (EC 4.1.3.12) (Fragment).
GN   Name=leuA;
OS   Buchnera aphidicola.
OG   Plasmid pBPp1.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Buchnera.
OX   NCBI_TaxID=9;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99028904; PubMed=9812361;
RA   Silva F.J., Van Ham R.C.H.J., Sabater B., Latorre A.;
RT   "Structure and evolution of the leucine plasmids carried by the
RT   endosymbiont (Buchnera aphidicola) from aphids of the family
RT   Aphididae.";
RL   FEMS Microbiol. Lett. 168:43-49(1998).
DR   EMBL; AJ006877; CAA07302.1; -.
DR   GO; GO:0016829; F:lyase activity; IEA.
KW   Lyase; Plasmid.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1538 MW;  1BD1D0320390C050 CRC64;

  Query Match             28.2%;  Score 20;  DB 2;  Length 13;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+04;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LNSKI 5
              :|||:
Db          1 MNSKV 5


RESULT 14
Q16118
```

```
ID   Q16118      PRELIMINARY;      PRT;    14 AA.
AC   Q16118;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Cyclic adenosine 3',5'-monophosphate response element binding protein
DE   (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94158910; PubMed=8114764;
RA   Waeber G., Meyer T.E., LeSieur M., Hermann H.L., Gerard N.,
RA   Habener J.H.;
RT   "Developmental stage-specific expression of cyclic adenosine 3',5'-
RT   monophosphate response element binding protein CREB during
RT   spermatogenesis involves alternative exon splicing.";
RL   Mol. Endocrinol. 7:1501-1501(1993).
DR   EMBL; S68577; AAB29985.2; -.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1698 MW;  C42ECDBEE6B7CEB6 CRC64;

  Query Match             28.2%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+04;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          8 KIVSQE 13
              ||| |:
Db          3 KIVKQD 8


RESULT 15
Q6SE52
ID   Q6SE52      PRELIMINARY;      PRT;    14 AA.
AC   Q6SE52;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Pgi (Fragment).
OS   Drosophila simulans (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7240;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=14762063; DOI=10.1101/gr.1329204;
RA   Halligan D.L., Eyre-Walker A., Andolfatto P., Keightley P.D.;
RT   "Patterns of evolutionary constraints in intronic and intergenic DNA
RT   of Drosophila.";
RL   Genome Res. 14:273-279(2004).
DR   EMBL; AY459549; AAR23007.1; -.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1456 MW;  2C83E49CCD8E7E37 CRC64;

  Query Match             28.2%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   57.1%;  Pred. No. 1.6e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LNSKIAF 7
              || : ||
Db          8 LNQEAAF 14


Search completed: February 22, 2005, 09:37:49
Job time : 53.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-3
Perfect score: 71
Sequence: 1 LNSKIAFKIVSQEPA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	71	100.0	15	2	AAW04843	Aaw04843 Self epit
2	71	100.0	15	2	AAW64815	Aaw64815 Desmoglei
3	71	100.0	15	2	AAW78814	Aaw78814 Desmoglei
4	71	100.0	15	3	AAB33625	Aab33625 MHC class
5	71	100.0	15	4	AAG93721	Aag93721 Human des
6	71	100.0	15	5	AAO17033	Aao17033 Desmoglei
7	71	100.0	15	6	ABU96577	Abu96577 MHC class
8	71	100.0	15	8	ADQ14318	Adq14318 Human des
9	71	100.0	15	8	ADR41700	Adr41700 Desmoglei
10	71	100.0	15	8	ADS14310	Ads14310 Desmoglei
11	31	43.7	15	5	ABB04739	Abb04739 Human hep
12	31	43.7	15	8	ADJ34233	Adj34233 Human sec
13	30	42.3	15	8	ADL26206	Adl26206 Synthetic
14	30	42.3	15	8	ADL26187	Adl26187 Synthetic
15	29	40.8	15	2	AAW85386	Aaw85386 Helper T-
16	29	40.8	15	2	AAW85174	Aaw85174 Helper T-
17	29	40.8	15	3	AAB13871	Aab13871 L2/HNK1 c
18	29	40.8	15	8	ADL26192	Adl26192 Synthetic
19	28	39.4	10	5	AAE24272	Aae24272 Murine E-
20	27	38.0	10	4	AAB73139	Aab73139 Tumour an
21	27	38.0	10	8	ADQ35117	Adq35117 Novel pep
22	27	38.0	10	8	ADQ35116	Adq35116 Novel pep
23	27	38.0	11	8	ADQ35123	Adq35123 Novel pep
24	27	38.0	15	5	ABP55461	Abp55461 Human zin
25	27	38.0	15	8	ADL26185	Adl26185 Synthetic
26	27	38.0	15	8	ADL26186	Adl26186 Synthetic
27	27	38.0	15	8	ADL26205	Adl26205 Synthetic
28	27	38.0	15	8	ADL26204	Adl26204 Synthetic
29	26	36.6	11	2	AAM53017	Aam53017 Antifunga
30	26	36.6	13	4	AAB72987	Aab72987 Human GAD
31	26	36.6	15	4	ABP24630	Abp24630 HIV DR su
32	26	36.6	15	4	AAJ03184	Aaj03184 Hepatitis
33	26	36.6	15	4	AAJ03491	Aaj03491 Hepatitis
34	26	36.6	15	8	ADL70833	Adl70833 PTP-Sap1
35	26	36.6	15	8	ADL26190	Adl26190 Synthetic
36	26	36.6	15	8	ADL26191	Adl26191 Synthetic
37	25.5	35.9	15	1	AAP82794	Aap82794 Von Wille
38	25.5	35.9	15	2	AAR40245	Aar40245 von Wille
39	25	35.2	9	4	ABP17290	Abp17290 HIV B27 s
40	25	35.2	10	4	ABP17353	Abp17353 HIV B27 s
41	25	35.2	10	4	ABP17354	Abp17354 HIV B27 s
42	25	35.2	11	4	ABP17402	Abp17402 HIV B27 s
43	25	35.2	11	4	ABP17403	Abp17403 HIV B27 s
44	25	35.2	11	4	ABP17401	Abp17401 HIV B27 s
45	25	35.2	12	2	AAW81200	Aaw81200 Synthetic

ALIGNMENTS

RESULT 1
AAW04843
ID AAW04843 standard; peptide; 15 AA.
XX
AC AAW04843;
XX
DT 18-FEB-1997 (first entry)
XX
DE Self epitope of desmoglein 3, implicated in autoimmune disease.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Homo sapiens.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 1; Page 39; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is derived from the human desmoglein

```
CC   3 protein (amino acids 190-204) and is implicated as a self epitope in
CC   pemphigus vulgaris. Peptides derived from the human desmoglein protein
CC   are described in AAW04841-47
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 71;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LNSKIAFKIVSQEPA 15
              |||||||||||||||
Db          1 LNSKIAFKIVSQEPA 15


RESULT 2
AAW64815
ID   AAW64815 standard; peptide; 15 AA.
XX
AC   AAW64815;
XX
DT   29-SEP-1998  (first entry)
XX
DE   Desmoglein-3 190-204.
XX
KW   Desmoglein; DG; gene therapy; pemphigus vulgaris; microparticle;
KW   autoantigen; autoimmune disease; MHC.
XX
OS   Homo sapiens.
XX
PN   US5783567-A.
XX
PD   21-JUL-1998.
XX
PF   22-JAN-1997;   97US-00787547.
XX
PR   22-JAN-1997;   97US-00787547.
XX
PA   (PANG-) PANGAEA PHARM INC.
XX
PI   Langer RS,  Hedley ML,  Curley JM;
XX
DR   WPI; 1998-427077/36.
XX
PT   Microparticle encapsulated nucleic acids - for recombinant expression of
PT   proteins e.g. in gene therapy.
XX
PS   Disclosure; Col 4; 42pp; English.
XX
CC   The patent describes a new preparation of microparticles each comprising
CC   a polymeric matrix and a nucleic acid. The polymeric matrix consists of
CC   one or more synthetic polymers having a solubility in water of less than
CC   1 mg/l (e.g. poly-lactic-co-glycolic acid); and at least 90% of the
CC   microparticles have a diameter of less than 100 microns. The
CC   microparticles are useful for the delivery of nucleic acids to phagocytic
CC   cells. In one embodiment the microparticles are less than 20 microns in
CC   diameter and the nucleic acid (preferably in closed circular form)
CC   includes an expression control sequence operatively linked to a coding
CC   sequence, where the expression product of the coding sequence is a
CC   polypeptide having a length and a sequence which permits it to bind to an
CC   MHC class I or II molecule. The expression product is thus an effective
CC   stimulator of an immune response in mammals. The present sequence, an
CC   antigenic portion of desmoglein 3, is an example of an MHC class II
CC   peptide which can be expressed by the nucleic acid. It is associated with
CC   pemphigus vulgaris
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 71;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LNSKIAFKIVSQEPA 15
              |||||||||||||||
Db          1 LNSKIAFKIVSQEPA 15


RESULT 3
AAW78814
ID   AAW78814 standard; peptide; 15 AA.
XX
AC   AAW78814;
XX
DT   17-NOV-1998  (first entry)
XX
DE   Desmoglein 3 protein fragment 190-204.
XX
KW   Microparticle; delivery; polymeric matrix; autoantigen; tumour antigen;
KW   class II associated peptide; pathogen; gene therapy; genetic disease;
KW   infection; downregulation; immune response.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9831398-A1.
XX
PD   23-JUL-1998.
XX
PF   22-JAN-1998;   98WO-US001499.
XX
PR   22-JAN-1997;   97US-00787547.
PR   06-JAN-1998;   98US-00003253.
XX
PA   (PANG-) PANGAEA PHARM INC.
XX
PI   Hedley ML,  Curley JM,  Langer RS,  Lunsford LB;
XX
DR   WPI; 1998-427556/36.
XX
PT   New preparations of microparticles - comprising a synthetic polymer
PT   matrix and nucleic acid comprising an expression vector for use in gene
PT   therapy.
XX
PS   Disclosure; Page 8; 101pp; English.
XX
CC   A microparticle preparation (MP) has been developed, consisting of
CC   microparticles having a diameter of less than 100 mu m. The MP comprises:
CC   (a) a polymeric matrix (PM) consisting of one or more synthetic polymers
CC   having a solubility in water of less that 1 mg/l; and (b) an expression
CC   vector selected from RNA molecules (at least 50% of which are closed
CC   circles) or circular plasmid DNA (at least 50% of which are supercoiled).
CC   Also described is a MP of at most 20 microns in diameter, comprising: (a)
CC   a PM; and (b) a NAM comprising an expression control sequence operatively
CC   linked to a coding sequence, where the coding sequence encodes an
CC   expression product selected from: (1) a polypeptide at least 7 amino
CC   acids in length, having a sequence identical to the sequence of: (i) a
CC   fragment of a naturally-occurring mammalian protein; or (ii) a fragment
CC   of a naturally-occurring protein from an infectious agent which infects a
CC   mammal; (2) a peptide having a length and sequence which permits it to
CC   bind to an MHC class I or II molecule; and (3) the polypeptide or the
CC   peptide linked to a trafficking sequence. AAW69763 to AAW69765, and
CC   AAW78793 to AAW78897 are peptide fragments for use in the present
CC   invention. The MPs are highly effective vehicles for the delivery of
CC   polynucleotides into phagocytic cells. They can be used for gene therapy,
CC   e.g. for treating genetic diseases, infections or tumours or for
CC   downregulating an immune response
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 71;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 2.3e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LNSKIAFKIVSQEPA 15
              |||||||||||||||
Db          1 LNSKIAFKIVSQEPA 15
```

RESULT 4
AAB33625
ID AAB33625 standard; peptide; 15 AA.
XX
AC AAB33625;
XX
DT 26-JAN-2001 (first entry)
XX
DE MHC class II associated immunogenic peptide SEQ ID 24.
XX
KW Microparticle; nucleic acid delivery; immunogenic peptide; MHC I; MHC II;
KW major histocompatibility complex; vaginal tissue; mucosal tissue.
XX
OS Unidentified.
XX
PN WO200053161-A2.
XX
PD 14-SEP-2000.
XX
PF 10-MAR-2000; 2000WO-US006578.
XX
PR 11-MAR-1999; 99US-00266463.
PR 27-MAY-1999; 99US-00321346.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Lunsford LB, Putnam D, Hedley ML;
XX
DR WPI; 2000-638130/61.
XX
PT Microparticles useful for administering a nucleic acid into the mucosal
PT tissue preferably vaginal tissue of an animal, comprises a polymeric
PT matrix, a lipid and a nucleic acid molecule.
XX
PS Claim 25; Page 11; 96pp; English.
XX
CC The present invention relates to microparticles which are less than 20
CC microns in diameter, which comprise a polymeric matrix, a lipid and a
CC nucleic acid molecule. The microparticle is specifically not encapsulated
CC in a liposome and does not comprise a cell. The nucleotide sequence
CC encodes an expression product that binds to major histocompatibility
CC complex (MHC) type I or II molecules. Peptides AAB33602-B33647 represent
CC MHC class II associated immunogenic peptides, and AAB33648-B33710
CC represent MHC class I associated immunogenic peptides. The peptides are
CC examples of the expression products of the nucleotide sequences which can
CC be included in the microparticles of the invention. Sequences AAB33711-
CC B33716 represent alternative expression products and nuclear localisation
CC signals also used in the invention. The microparticles are useful for
CC administering a nucleic acid into the mucosal tissue preferably vaginal
CC tissue of an animal
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNSKIAFKIVSQEPA 15
||||||||||||||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 5
AAG93721
ID AAG93721 standard; peptide; 15 AA.
XX
AC AAG93721;
XX
DT 17-SEP-2001 (first entry)
XX
DE Human desmoglein 3 peptide 1.
XX
KW Continuous flow production; microparticle; gene therapy;
KW antisense therapy; vaccination; treatment; autoimmune disease;
KW immune response modulation.
XX
OS Homo sapiens.
XX
PN WO200136583-A1.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-US031770.
XX
PR 19-NOV-1999; 99US-00443654.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Hsu Y, Tyo M;
XX
DR WPI; 2001-425203/45.
XX
PT Continuous production of microparticles containing nucleic acid for e.g.
PT gene therapy, comprises mixing a solution of polymeric material and
PT nucleic acid with a surfactant solution, removing solvent and drying.
XX
PS Disclosure; Page 9; 47pp; English.
XX
CC The present sequence is that of a peptide of the invention. The invention
CC relates to a method for scalable, continuous flow production of a nucleic
CC acid containing microparticle that maintains the structural integrity of
CC the associated nucleic acid and results in a microparticle having purity
CC suitable for introduction into an animal host. Microparticles prepared
CC according to the method can be used for delivery of a nucleic acid for
CC gene therapy, antisense therapy, vaccination, treatment of autoimmune
CC disease and either specific or non-specific modulation of an immune
CC response. The microparticles may also be used to deliver nucleic acid
CC encoding a protein or peptide useful in any kind of therapy. The method
CC is economical, aseptic and scalable. The method also enables control over
CC the size of microparticles. The microparticles produced are free of
CC impurities such as organic solvents and are readily dispersed in a wide
CC range of dispersing agents
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNSKIAFKIVSQEPA 15
||||||||||||||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 6
AAO17033
ID AAO17033 standard; peptide; 15 AA.
XX
AC AAO17033;
XX
DT 29-MAY-2002 (first entry)
XX
DE Desmoglein 3 residues 190-204 SEQ ID NO: 30.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX

PN WO200206316-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022263.
XX
PR 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX
DR WPI; 2002-195801/25.
XX
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
PS Disclosure; Page 26; 89pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a peptide described in the exemplification of the
CC invention
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNSKIAFKIVSQEPA 15
||||||||||||||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 7
ABU96577
ID ABU96577 standard; peptide; 15 AA.
XX
AC ABU96577;
XX
DT 12-AUG-2003 (first entry)
XX
DE MHC class II associated desmoglein 3 peptide 190-204.
XX
KW Microparticle; microsphere; polynucleotide delivery; phagocytic cell;
KW tumour; viral infection; bacterial infection; fungal infection;
KW protozoan infection; gene therapy; major histocompatability complex;
KW MHC class II.
XX
OS Unidentified.
XX
PN US2002182258-A1.
XX
PD 05-DEC-2002.
XX
PF 18-JUL-2001; 2001US-00909460.
XX
PR 22-JAN-1997; 97US-0035983P.
PR 06-JAN-1998; 98US-00003253.
PR 22-JAN-1998; 98WO-US001499.
PR 11-MAR-1999; 99US-00266463.
PR 27-MAY-1999; 99US-00321346.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Lunsford LB, Putnam D, Hedley ML;
XX
DR WPI; 2003-438782/41.
XX
PT Microparticles, useful as vehicles for delivery of polynucleotides to
PT phagocytic cells, comprises polymeric matrix, lipid, and nucleic acid
PT molecule.
XX
PS Disclosure; Page 3; 37pp; English.
XX
CC The invention relates to a microparticle (microsphere) less than 20
CC microns in diameter that comprises: (1) a polymeric matrix; (2) a lipid;
CC and (3) a nucleic acid molecule. The microparticle is not encapsulated in
CC a liposome and the microparticle does not comprise a cell. The
CC microparticles are used as vehicles for the delivery of polynucleotides
CC into phagocytic cells. The microparticles can be used to express antigens
CC to treat tumour cells or viral, bacterial, fungal or protozoan
CC infections. The microparticles can be made without adversely affecting
CC nucleic acid integrity. The present sequence represents the amino acid
CC sequence of a major histocompatability complex, MHC, class II associated
CC peptide
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNSKIAFKIVSQEPA 15
||||||||||||||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 8
ADQ14318
ID ADQ14318 standard; peptide; 15 AA.
XX
AC ADQ14318;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human desmoglein 3 (DSG3) antigenic determinant (residues 190-204).
XX
KW Notch signalling; modulator; antigen; antigenic determinant;
KW immunomodulator; immune disorder; immune response; immune tolerance;
KW peripheral T-cell activation; regulatory T-cell; T reg; tumour; cancer;
KW autoimmune disorder; allergy; transplant rejection; immunosuppressive;
KW cytostatic; antiallergic; vaccine; human; autoantigen; desmoglein 3;
KW DSG3; pemphigus; MHC class II peptide; major histocompatibility complex;
KW antigenic determinant; epitope.
XX
OS Homo sapiens.
XX
PN WO2004060262-A2.
XX
PD 22-JUL-2004.
XX
PF 07-JAN-2004; 2004WO-GB000046.
XX
PR 07-JAN-2003; 2003GB-00000234.
PR 23-JAN-2003; 2003GB-00001510.
PR 23-JAN-2003; 2003GB-00001512.
PR 23-JAN-2003; 2003GB-00001513.
PR 23-JAN-2003; 2003GB-00001515.
PR 23-JAN-2003; 2003GB-00001518.
PR 23-JAN-2003; 2003GB-00001519.
PR 23-JAN-2003; 2003GB-00001521.
PR 23-JAN-2003; 2003GB-00001522.
PR 23-JAN-2003; 2003GB-00001524.

PR 23-JAN-2003; 2003GB-00001526.
PR 23-JAN-2003; 2003GB-00001527.
PR 23-JAN-2003; 2003GB-00001529.
PR 22-MAR-2003; 2003GB-00006621.
PR 04-APR-2003; 2003WO-GB001525.
PR 24-MAY-2003; 2003GB-00012062.
PR 01-AUG-2003; 2003WO-GB003285.
PR 03-OCT-2003; 2003GB-00023130.
XX
PA (LORA-) LORANTIS LTD.
XX
PI Bodmer MW, Briend ECP, Champion BR, Lennard AC, Mckenzie GJ;
PI Tugal T, Ward GA, Young LL;
XX
DR WPI; 2004-534298/51.
XX
PT New product for modulating the immune system, comprises a pharmaceutical
PT support matrix bearing modulators of Notch signaling, and an antigen or
PT antigenic determinant, or a polynucleotide coding for the antigen or
PT determinant.
XX
PS Disclosure; Page 126; 294pp; English.
XX
CC The invention relates to a product comprising (1) a pharmaceutical
CC support matrix for in vivo administration bearing modulators of Notch
CC signalling (especially a Delta or Serrate/Jagged protein or fragment or
CC homologue thereof); and (2) an antigen or antigenic determinant, or
CC polynucleotide encoding an antigen or antigenic determinant. The product
CC acts as a combined preparation for modulation of the immune system or for
CC modulation of an immune response to the antigen or antigenic determinant.
CC The invention also relates to a pharmaceutical composition comprising the
CC product; a particle with a maximum linear dimension of less than 500
CC (preferably 30-70) nm having several bound modulators of Notch signalling
CC ; a method of modulating Notch signalling; methods of treating an immune
CC disorder, for reducing an immune response, for promoting immune tolerance
CC in a mammal; and a method for increasing an immune response to a tumour
CC or pathogen antigen or its antigenic determinant in a mammal. The
CC composition and methods are useful for modulating peripheral T-cell
CC activation, for generating regulatory T-cells (T regs), for reducing an
CC immune response to an antigen or antigenic determinant, for promoting
CC immune tolerance to an antigen or antigenic determinant, or for treating
CC tumours, autoimmune disease, allergies or transplant rejection. The
CC present sequence represents an antigenic determinant derived from a human
CC autoantigen which may be used in a product of the invention.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNSKIAFKIVSQEPA 15
||||||||||||||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 9
ADR41700
ID ADR41700 standard; peptide; 15 AA.
XX
AC ADR41700;
XX
DT 21-OCT-2004 (first entry)
XX
DE Desmoglein 3 (DSG3) (aa 190-204), class II MHC associated autoantigen.
XX
KW DSG3; Desmoglein 3; autoantigen; notch signalling pathway;
KW autoimmune disorder; bystander effect; suppression; DSL domain;
KW EGF domain; Goodpasture's disease; Wegener's granulomatosis; anaemia;
KW thrombocytopenia; gastritis; hepatitis; vasculitis; scleroderma;
KW myositis; arthritis; Systemic Lupus Erythematosus; SLE;
KW Sjogren's syndrome; hepatic fibrosis; liver cirrhosis; thyroiditis;
KW dermatitis; placental dysfunction; eclampsia;
KW inflammatory related gynaecological disease; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; Huntington disease;
KW encephalitis; psychiatric disorder; Down's syndrome; stroke; exogenous;
KW bystander antigen; multiple sclerosis; delta serrate lag;
KW inflammatory bowel disease; notch receptor.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO2004064863-A1.
XX
PD 05-AUG-2004.
XX
PF 23-JAN-2004; 2004WO-GB000263.
XX
PR 23-JAN-2003; 2003GB-00001510.
PR 23-JAN-2003; 2003GB-00001512.
PR 23-JAN-2003; 2003GB-00001513.
PR 23-JAN-2003; 2003GB-00001515.
PR 23-JAN-2003; 2003GB-00001518.
PR 23-JAN-2003; 2003GB-00001519.
PR 23-JAN-2003; 2003GB-00001521.
PR 23-JAN-2003; 2003GB-00001522.
PR 23-JAN-2003; 2003GB-00001524.
PR 23-JAN-2003; 2003GB-00001526.
PR 23-JAN-2003; 2003GB-00001527.
PR 23-JAN-2003; 2003GB-00001529.
PR 04-APR-2003; 2003WO-GB001525.
PR 24-MAY-2003; 2003GB-00012062.
PR 01-AUG-2003; 2003WO-GB003285.
PR 03-OCT-2003; 2003GB-00023130.
PR 07-JAN-2004; 2004WO-GB000046.
XX
PA (LORA-) LORANTIS LTD.
XX
PI Champion BR, Ragno S, Young LL;
XX
DR WPI; 2004-562091/54.
XX
PT New product having a modulator of the Notch signaling pathway, useful for
PT modulating an immune response in autoimmune disorders, such as anemia,
PT gastritis, hepatitis, scleroderma and myositis.
XX
PS Disclosure; Page 66; 244pp; English.
XX
CC The invention relates to the modulation of immune function through a
CC notch signalling pathway for the prevention of autoimmune diseases. It
CC has been found that the notch signalling pathway provides a bystander
CC effect or bystander suppression effect, which can be used in a wide
CC variety of ways to suppress unwanted immune responses in immune diseases
CC and disorders. Autoimmune diseases are characterised by immune responses
CC that are directed against self antigens. T lymphocytes are activated upon
CC recognition of a self antigen and/or a foreign antigen as a complex with
CC self major histocompatibility complex (MHC) gene products on the surface
CC of antigen presenting cells (APC). The invention provides the method of
CC modulating of an immune response, modulator information and a
CC pharmaceutical kit for suppression of an immune response. The modulator
CC of the notch signalling pathway is an agent which activates the notch
CC receptor or a polynucleotide which codes for such an agent. It comprises
CC a protein or polypeptide comprising a notch ligand DSL (delta serrate
CC lag) domain, notch ligand EGF domain, optionally all or part of a notch
CC ligand N terminal domain, and optionally one or more heterologous amino
CC acid or a polynucleotide sequences. The modulator can be a fusion protein
CC comprising a segment of a notch ligand extracellular domain and an
CC immunoglobulin Fc segment. The disorders include Goodpasture's disease,
CC Wegener's granulomatosis, autoimmune anaemia, thrombocytopenia,
CC gastritis, autoimmune hepatitis, inflammatory bowel disease, autoimmune
CC vasculitis, scleroderma, myositis, autoimmune arthritis, Systemic Lupus
CC Erythematosus (SLE) or Sjogren's syndrome, hepatic fibrosis, liver
CC cirrhosis, thyroiditis, dermatitis, placental dysfunction, eclampsia,
CC inflammatory related gynaecological diseases, neurodegenerative disorders
CC (such as Alzheimer's disease, Parkinson's disease, Huntington disease)
CC encephalitis, psychiatric disorders, Down's syndrome, stroke, multiple

CC sclerosis, etc. The invention discloses a method for generating immune
CC suppression at a disease locus by administering an exogenous antigen. It
CC also provides the use of modulator or activator of notch signalling in
CC simultaneous, separate or sequential combination with a bystander antigen
CC or antigenic determinant for reducing an immune response to a target
CC antigen. The presented protein sequence is the desmoglein 3 (DSG3) (aa
CC 190-204), class II MHC associated autoantigen.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNSKIAFKIVSQEPA 15
||||||||||||||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 10
ADS14310
ID ADS14310 standard; peptide; 15 AA.
XX
AC ADS14310;
XX
DT 16-DEC-2004 (first entry)
XX
DE Desmoglein 3 antigenic peptide residues 190-204.
XX
KW Cytostatic; Immunosuppressive; Antidiabetic; Neuroprotective;
KW Antiarthritic; Antirheumatic; Antiallergic; Vaccine; Notch signaling;
KW Notch; Notch ligand; Delta protein; Serrate protein; Jagged protein;
KW multiple sclerosis; rheumatoid arthritis; diabetes; allergy;
KW immune disorder; autoimmune disease; graft rejection; cancer;
KW organ transplant; desmoglein 3.
XX
OS Unidentified.
XX
PN WO2004083372-A2.
XX
PD 30-SEP-2004.
XX
PF 22-MAR-2004; 2004WO-GB001229.
XX
PR 21-MAR-2003; 2003GB-00006582.
PR 21-MAR-2003; 2003GB-00006583.
PR 22-MAR-2003; 2003GB-00006621.
PR 22-MAR-2003; 2003GB-00006622.
PR 22-MAR-2003; 2003GB-00006624.
PR 22-MAR-2003; 2003GB-00006626.
PR 22-MAR-2003; 2003GB-00006640.
PR 22-MAR-2003; 2003GB-00006644.
PR 22-MAR-2003; 2003GB-00006650.
PR 22-MAR-2003; 2003GB-00006651.
PR 22-MAR-2003; 2003GB-00006654.
XX
PA (LORA-) LORANTIS LTD.
XX
PI Champion BR, Ragno S;
XX
DR WPI; 2004-709927/69.
XX
PT Particle capable of being inserted into or taken up by cell useful for
PT modulating immune response to antigen in subject, comprises
PT polynucleotide coding for modulator of Notch signaling, and
PT polynucleotide coding for antigen.
XX
PS Disclosure; Page 118; 278pp; English.
XX
CC The present invention relates to a particle (I) capable of being inserted
CC into or taken up by a cell comprising (i) a polynucleotide coding for a
CC modulator of Notch signaling, and (ii) a polynucleotide coding for an
CC antigen or antigenic determinant. The first polynucleotide sequence codes
CC for a Notch ligand such as a Delta or Serrate/Jagged protein or its
CC fragment, derivative, homologue, analogue or allelic variant, or for a
CC protein which comprises a Notch ligand DSL domain and at least one Notch
CC ligand EGF-like domain and optionally a membrane binding or transmembrane
CC domain. The first and second sequences are operably linked to one or more
CC promoters or enhancers or polyadenylation sequences. The antigen or
CC antigenic determinant is an allergen, autoantigen, Major
CC Histocompatibility complex (MHC) (transplant) antigen, pathogen antigen,
CC tumour antigen or their antigenic determinant. (I) is useful for
CC modulating an immune response to an antigen in a subject. Pharmaceutical
CC compositions comprising (I) are useful for treating or preventing
CC conditions mediated by T cells, such as multiple sclerosis, rheumatoid
CC arthritis, diabetes, allergy, for treating immune disorders such as
CC autoimmune diseases of graft rejection such as allograft rejection,
CC treating cancer and organ transplants. The present sequence is a
CC desmoglein 3 antigenic peptide (a class II MHC-associated autoantigen
CC peptide), which can be used as an antigen to prepare the particle of the
CC invention.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 71; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNSKIAFKIVSQEPA 15
||||||||||||||||
Db 1 LNSKIAFKIVSQEPA 15

RESULT 11
ABB04739
ID ABB04739 standard; peptide; 15 AA.
XX
AC ABB04739;
XX
DT 11-MAR-2002 (first entry)
XX
DE Human heparan sulfate 3-O-sulfotransferase 19 peptide SEQ ID NO:7.
XX
KW Human; heparan sulfate 3-O-sulfotransferase 19; haemorrhagic disease;
KW thrombus embolism; haemopathy; myocardial infarction; tumour;
KW inflammation; immunological disease; HIV infection.
XX
OS Homo sapiens.
XX
PN CN1311305-A.
XX
PD 05-SEP-2001.
XX
PF 02-MAR-2000; 2000CN-00111794.
XX
PR 02-MAR-2000; 2000CN-00111794.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-049921/07.
XX
PT New polypeptide-heparan sulfate 3-0-sulfotransferase 19 and
PT polynucleotide for coding such polypeptide.
XX
PS Example 6; Page 19 (Disclosure); 34pp; Chinese.
XX
CC The present invention describes human heparan sulfate 3-O-
CC sulfotransferase 19 protein (I). The present invention also describes a
CC method of applying (I) in the treatment of various diseases, such as
CC haemorrhagic diseases, thrombus embolism, other haemopathy, myocardial
CC infarction, various tumours, inflammation, immunological diseases and HIV
CC infection. The present invention also describes the agonist resisting (I)
CC and its treatment effect. The present sequence represents the N-terminal
CC peptide of (I) which is used in an example from the present invention

```
XX
SQ   Sequence 15 AA;

  Query Match             43.7%;  Score 31;  DB 5;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 96;
  Matches    6;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 LNSKIAFKIVSQEP 14
              :|| |   |: :||
Db          1 MNSSIKLLIIVREP 14


RESULT 12
ADJ34233
ID   ADJ34233 standard; protein; 15 AA.
XX
AC   ADJ34233;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human secreted protein NOV9e.
XX
KW   Human; NOVX; secreted protein; cancer; diabetes; obesity;
KW   endocrine disorder; CNS disorder; inflammatory disorder; gene therapy.
XX
OS   Homo sapiens.
XX
PN   WO2004000997-A2.
XX
PD   31-DEC-2003.
XX
PF   04-JUN-2003; 2003WO-US017512.
XX
PR   19-MAR-2002; 2002US-0365491P.
PR   04-JUN-2002; 2002US-0385504P.
PR   05-JUN-2002; 2002US-0386041P.
PR   06-JUN-2002; 2002US-0386453P.
PR   06-JUN-2002; 2002US-0386974P.
PR   07-JUN-2002; 2002US-0386816P.
PR   07-JUN-2002; 2002US-0387002P.
PR   10-JUN-2002; 2002US-0387540P.
PR   11-JUN-2002; 2002US-0387659P.
PR   12-JUN-2002; 2002US-0387934P.
PR   13-JUN-2002; 2002US-0389123P.
PR   17-JUN-2002; 2002US-0389729P.
PR   17-JUN-2002; 2002US-0389742P.
PR   19-JUN-2002; 2002US-0390006P.
PR   17-JUL-2002; 2002US-0396706P.
PR   12-AUG-2002; 2002US-0402832P.
PR   13-AUG-2002; 2002US-0403486P.
PR   14-AUG-2002; 2002US-0403522P.
PR   15-AUG-2002; 2002US-0403748P.
PR   06-NOV-2002; 2002US-0387037P.
PR   03-JUN-2003; 2003US-00454246.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Anderson DW,  Boldog FL,  Burgess CE,  Casman SJ,  Edinger SR;
PI   Eisen A,  Ellerman K,  Gerlach VL,  Gorman L,  Guo X,  Gusev VY,  Ji W;
PI   Li L,  Macdougall JR,  Malyankar UM,  Millet I,  Ort T,  Padigaru M;
PI   Prayaga SK,  Patturajan M,  Pena CEA,  Peyman JA,  Rieger DK;
PI   Rothenberg ME,  Sciore P,  Shenoy SG,  Smithson G,  Spytek KA,  Stone DJ;
PI   Taupier RJ,  Tchernev VT,  Vernet CAM,  Voss EZ,  Zerhusen BD,  Zhong M;
XX
DR   WPI; 2004-082483/08.
DR   N-PSDB; ADJ34232.
XX
PT   New isolated NOVX polypeptides useful for treating, preventing and
PT   diagnosing pathological conditions with NOVX-associated disorders, such
PT   as cancer, obesity, diabetes and inflammatory or CNS diseases.
XX
PS   Claim 1; SEQ ID NO 122; 418pp; English.
XX
CC   The invention relates to a new isolated polypeptide (designated NOVX)
CC   comprising one of 141 fully defined sequences, their mature forms, a
CC   protein comprising one or more conservative substitutions or having at
CC   least 95% identity to one of the 141 proteins. Also included are a
CC   composition comprising NOVX (or a NOVX nucleic acid molecule (NA)), a kit
CC   comprising the composition of NOVX in one or more containers, an isolated
CC   nucleic acid molecule encoding a NOVX protein, producing NOVX (comprising
CC   culturing a cell under conditions that lead to expression of the
CC   polypeptide, where the cell comprises a vector comprising NOVX NA),
CC   identifying an agent that binds to NOVX, identifying a potential
CC   therapeutic agent for use in the treatment of a pathology that is related
CC   to aberrant expression or physiological interactions of NOVX, screening
CC   for a modulator of activity of or latency or predisposition to a
CC   pathology associated with NOVX, modulating the activity of NOVX, treating
CC   or preventing a pathology associated with NOVX, treating a pathological
CC   state in a mammal, a vector comprising the NOVX nucleic acid molecule, a
CC   cell comprising the vector, an antibody that immunospecifically binds to
CC   NOVX, determining the presence or amount of NOVX or the nucleic acid
CC   molecule in a sample, and determining the presence of or predisposition
CC   to a disease associated with altered levels of expression of NOVX or the
CC   nucleic acid molecule in a first mammalian subject. The methods and
CC   compositions of the present invention are useful for the diagnosis and
CC   treatment of disorders associated with aberrant expression or activity of
CC   the NOVX polypeptide, such as cancer, diabetes, obesity, and endocrine,
CC   CNS and inflammatory disorders. They can also be used in various
CC   detection and screening assays, chromosome mapping, tissue typing, gene
CC   therapy and predictive medicine. The present sequence represents a NOVX
CC   protein.
XX
SQ   Sequence 15 AA;

  Query Match             43.7%;  Score 31;  DB 8;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 96;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NSKIAFKIV 10
              |||:||  |
Db          2 NSKVAFSAV 10


RESULT 13
ADL26206
ID   ADL26206 standard; peptide; 15 AA.
XX
AC   ADL26206;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Synthetic peptide B121 derived from a conserved region of HCV.
XX
KW   HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;
KW   major histocompatibility complex; human leukocyte antigen.
XX
OS   Synthetic.
XX
PN   WO2004024182-A2.
XX
PD   25-MAR-2004.
XX
PF   27-AUG-2003; 2003WO-EP009482.
XX
PR   13-SEP-2002; 2002AT-00001376.
PR   27-FEB-2003; 2003WO-EP002005.
PR   11-JUL-2003; 2003EP-00450171.
XX
PA   (INTE-) INTERCELL AG.
XX
PI   Buschle M,  Habel A,  Klade C,  Mattner F,  Otava O,  Vytvytska O;
PI   Zauner W,  Zinke S,  Kirlappos H;
XX
DR   WPI; 2004-269899/25.
XX
PT   Isolating Hepatitis C Virus peptides (HPs) which have a binding capacity
```

PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
PT molecule by separating the complex from the HCV-peptides which do not
PT bind to the molecule.
XX
PS Example 1; Page 31; 73pp; English.
XX
CC The invention relates to a novel method for isolating Hepatitis C Virus
CC (HCV) peptides (HPs). The method of the invention has virucide activity,
CC and may be useful in producing a vaccine. The method is useful for
CC isolating Hepatitis C Virus peptides (HPs) which have a binding capacity
CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T
CC cells, a T cell clone or a T cell population or preparation is useful for
CC identifying heteroclitic epitopes or for preparing a composition for
CC treating HCV infection. The present sequence represents a synthetic
CC peptide derived from a conserved region of HCV.
XX
SQ Sequence 15 AA;

Query Match 42.3%; Score 30; DB 8; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 IAFKIVSQE 13
:||||:| |
Db 7 VAFKIMSGE 15

RESULT 14
ADL26187
ID ADL26187 standard; peptide; 15 AA.
XX
AC ADL26187;
XX
DT 17-JUN-2004 (first entry)
XX
DE Synthetic peptide B102 derived from a conserved region of HCV.
XX
KW HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;
KW major histocompatibility complex; human leukocyte antigen.
XX
OS Synthetic.
XX
PN WO2004024182-A2.
XX
PD 25-MAR-2004.
XX
PF 27-AUG-2003; 2003WO-EP009482.
XX
PR 13-SEP-2002; 2002AT-00001376.
PR 27-FEB-2003; 2003WO-EP002005.
PR 11-JUL-2003; 2003EP-00450171.
XX
PA (INTE-) INTERCELL AG.
XX
PI Buschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;
PI Zauner W, Zinke S, Kirlappos H;
XX
DR WPI; 2004-269899/25.
XX
PT Isolating Hepatitis C Virus peptides (HPs) which have a binding capacity
PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
PT molecule by separating the complex from the HCV-peptides which do not
PT bind to the molecule.
XX
PS Example 1; Page 31; 73pp; English.
XX
CC The invention relates to a novel method for isolating Hepatitis C Virus
CC (HCV) peptides (HPs). The method of the invention has virucide activity,
CC and may be useful in producing a vaccine. The method is useful for
CC isolating Hepatitis C Virus peptides (HPs) which have a binding capacity
CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T
CC cells, a T cell clone or a T cell population or preparation is useful for
CC identifying heteroclitic epitopes or for preparing a composition for
CC treating HCV infection. The present sequence represents a synthetic
CC peptide derived from a conserved region of HCV.
XX
SQ Sequence 15 AA;

Query Match 42.3%; Score 30; DB 8; Length 15;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 IAFKIVSQE 13
:||||:| |
Db 7 VAFKIMSGE 15

RESULT 15
AAW85386
ID AAW85386 standard; peptide; 15 AA.
XX
AC AAW85386;
XX
DT 16-FEB-1999 (first entry)
XX
DE Helper T-cell class II peptide derived from EXP1 protein.
XX
KW Helper T-cell peptide; human leucocyte antigen; HLA; DR4w4; DR1; DR7;
KW cytotoxic T lymphocyte; CTL; hepatitis; autoimmune disease;
KW acquired immune deficiency syndrome; malaria; cancer;
KW allograft rejection; allergy; Lyme disease; hepatitis;
KW post-streptococcal endocarditis; glomerulonephritis;
KW food hypersensitivity.
XX
OS Synthetic.
OS Plasmodium falciparum.
XX
PN WO9832456-A1.
XX
PD 30-JUL-1998.
XX
PF 23-JAN-1998; 98WO-US001373.
XX
PR 23-JAN-1997; 97US-0036713P.
PR 07-FEB-1997; 97US-0037432P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S;
XX
DR WPI; 1998-427679/36.
XX
PT Composition containing peptide that induces cytotoxic T lymphocyte
PT response, and helper peptide - can bind to human leucocyte antigen
PT alleles, used to treat or prevent cancers, parasitic infections and
PT autoimmune disease.
XX
PS Disclosure; Page 41; 51pp; English.
XX
CC AAW85284-451 represent helper T-cell class II peptides, which can bind to
CC the human leucocyte antigens (HLA) DR4w4, DR1 and DR7. The peptides are
CC used in the course of the invention. The specification describes peptides
CC that that induce a cytotoxic T lymphocyte (CTL) response, and T-helper
CC peptides, that are used together to generate a CTL response for the
CC treatment or prevention of viral, fungal, bacterial or parasitic
CC infections (e.g. hepatitis, acquired immune deficiency syndrome or
CC malaria) or cancer (e.g. renal or cervical carcinoma, lymphoma, prostate
CC cancer or condyloma acuminatum). Helper T-cell peptides may be used alone
CC to induce a helper T cell response, e.g. in cases of autoimmune disease,
CC allograft rejection, allergy, Lyme disease, hepatitis, post-streptococcal
CC endocarditis, glomerulonephritis and food hypersensitivity
XX
SQ Sequence 15 AA;

Query Match 40.8%; Score 29; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          7 FKIVSQEPA 15
              ||| | :||
Db          4 FKIGSSDPA 12
```

Search completed: February 22, 2005, 09:24:35
Job time : 67.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-4
Perfect score: 76
Sequence: 1 TPMFLLSRNTGEVRT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	31.6	11	2	PH0919	T-cell receptor be
2	23	30.3	14	2	S22236	lipoxygenase (EC 1
3	23	30.3	14	2	PH1705	Ig heavy chain V r
4	23	30.3	15	2	C43334	orf3 3' to aadR -
5	23	30.3	15	2	PA0075	fructose-bisphosph
6	22	28.9	15	2	PA0102	fructose-bisphosph
7	21	27.6	11	2	S33782	acetolactate synth
8	21	27.6	14	2	PH0915	T-cell receptor be
9	20	26.3	6	2	A27696	contraction-inhibi
10	20	26.3	8	2	PT0691	T-cell receptor be
11	20	26.3	11	2	B41835	translation elonga
12	20	26.3	12	2	G49410	t-complex polypept
13	20	26.3	15	2	PH0750	T-cell receptor be
14	20	26.3	15	2	PC4213	bphB protein - Com
15	19	25.0	6	2	B27696	contraction-inhibi
16	19	25.0	10	2	A49581	sialokinin I - yel
17	19	25.0	11	2	S51732	T-cell receptor al
18	19	25.0	12	2	PT0216	T-cell receptor be
19	19	25.0	12	2	PH0920	T-cell receptor be
20	19	25.0	13	2	S47390	T-cell antigen rec
21	19	25.0	13	2	PH0796	T-cell receptor al
22	19	25.0	13	2	PH0799	T-cell receptor al
23	19	25.0	13	2	PH0783	T-cell receptor al
24	19	25.0	15	2	S47367	T-cell antigen rec
25	18	23.7	6	2	PT0518	T-cell receptor be
26	18	23.7	10	2	PH0926	T-cell receptor be
27	18	23.7	12	2	S49547	hypothetical prote
28	18	23.7	13	2	B36042	oxfX protein - Esc
29	18	23.7	13	2	A26999	carboxylesterase (
30	18	23.7	13	2	I77387	AMP deaminase - ra
31	18	23.7	14	2	S58426	spermadhesin AWN h
32	18	23.7	14	2	B56819	PS I complex subun
33	18	23.7	15	2	I58116	Dp116 - human
34	18	23.7	15	2	E91061	hypothetical prote
35	18	23.7	15	2	A26228	spot 42 protein -
36	18	23.7	15	2	A40634	orf19 3' of eryK -
37	18	23.7	15	2	S03955	acidic fibroblast
38	18	23.7	15	4	I52698	hypothetical THRA1
39	17.5	23.0	14	2	D45474	thrombospondin 2 -
40	17	22.4	9	2	A31576	xylose isomerase (
41	17	22.4	9	2	G85802	hypothetical prote
42	17	22.4	10	2	S65387	cytochrome-c oxida
43	17	22.4	11	2	A54348	N-acetylglucosamin
44	17	22.4	11	2	S60354	retinal oxidase -
45	17	22.4	11	2	PT0211	T-cell receptor al

ALIGNMENTS

RESULT 1
PH0919
T-cell receptor beta chain V-D-J region (isolate 5) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C;Accession: PH0919
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic
A;Reference number: PH0891; MUID:92078857; PMID:1836012
A;Accession: PH0919
A;Molecule type: mRNA
A;Residues: 1-11 <GOL>
A;Experimental source: concanavalin A-activated lymphoblast
A;Note: the authors translated the codon CAG for residue 11 as Glu
C;Keywords: T-cell receptor

Query Match 31.6%; Score 24; DB 2; Length 11;
Best Local Similarity 80.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          8 RNTGE 12
              ||||:
Db          7 RNTGQ 11
```

RESULT 2
S22236
lipoxygenase (EC 1.13.11.12) 1 - barley (fragment)
C;Species: Hordeum vulgare (barley)
C;Date: 19-Mar-1997 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C;Accession: S22236
R;Doderer, A.; Kokkelink, I.; van der Veen, S.; Valk, B.E.; Schram, A.W.; Douma, A.C.
Biochim. Biophys. Acta 1120, 97-104, 1992
A;Title: Purification and characterization of two lipoxygenase isoenzymes from germinatir
A;Reference number: S21772; MUID:92207997; PMID:1554746
A;Accession: S22236
A;Molecule type: protein
A;Residues: 1-14 <DOD>
A;Experimental source: var. Triumph, seed
C;Function:
A;Description: catalyzes the peroxidation of polyunsaturates fatty acids to their corresp
C;Superfamily: lipoxygenase
C;Keywords: monomer; oxidoreductase; seed

Query Match 30.3%; Score 23; DB 2; Length 14;
Best Local Similarity 45.5%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```
Qy          2 PMFLLSRNTGE 12
              |  ||  || :
Db          2 PYMLLYPNTSD 12
```

```
RESULT 3
PH1705
Ig heavy chain V region (clone ASC-1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C;Accession: PH1705
R;McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A;Title: Antigen-driven B cell differentiation in vivo.
A;Reference number: PH1675; MUID:93301607; PMID:8315385
A;Accession: PH1705
A;Molecule type: mRNA
A;Residues: 1-14 <MCH>
A;Experimental source: B cell
A;Note: the authors translated the codon GTA for residue 11 as Thr and ACA for residue 1
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

  Query Match             30.3%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   62.5%;  Pred. No. 7.5e+02;
  Matches    5;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          8 RNTGEVRT 15
              |||  | |
Db          6 RNTKSVET 13


RESULT 4
C43334
orf3 3' to aadR - Rhodopseudomonas palustris (fragment)
C;Species: Rhodopseudomonas palustris
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: C43334
R;Dispensa, M.; Thomas, C.T.; Kim, M.K.; Perrotta, J.A.; Gibson, J.; Harwood, C.S.
J. Bacteriol. 174, 5803-5813, 1992
A;Title: Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate is dependen
A;Reference number: A43334; MUID:92394882; PMID:1522059
A;Accession: C43334
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-15 <DIS>
A;Cross-references: UNIPROT:Q02006; GB:M92426; NID:g151870; PIDN:AAA26091.1; PID:g551951
A;Note: sequence extracted from NCBI backbone (NCBIN:112964, NCBIP:112967)

  Query Match             30.3%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TPMFLLSRN 9
              || :| :|:
Db          5 TPAYLRARH 13


RESULT 5
PA0075
fructose-bisphosphate aldolase (EC 4.1.2.13) I - fungus (Fusarium sporotrichioides) (fra
N;Alternate names: aldolase; fructose-1,6,-biophosphate triosephosphate-lase
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0075; PA0077
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0075
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Z5
A;Note: this form (I) had a molecular weight of 30.6K and an isoelectric point of 5.3
A;Accession: PA0077
A;Molecule type: protein
A;Residues: 1-15 <CH2>
A;Note: this form (II) had a molecular weight of 31.6K and an isoelectric point of 5.4
C;Keywords: aldehyde-lyase; carbon-carbon lyase

  Query Match             30.3%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 LLSRNTGEV 13
              :|||::| :
Db          5 VLSRDSGVI 13


RESULT 6
PA0102
fructose-bisphosphate aldolase (EC 4.1.2.13) III - fungus (Fusarium sporotrichioides) (fr
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0102
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrichi
A;Reference number: PA0051
A;Accession: PA0102
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Z4
C;Keywords: aldehyde-lyase; carbon-carbon lyase

  Query Match             28.9%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 LLSRNTGEV 13
              :||| :| :
Db          5 VLSRKSGVI 13


RESULT 7
S33782
acetolactate synthase (EC 4.1.3.18) small chain, valine-sensitive - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 19-Mar-1997 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C;Accession: S33782
R;Yang, J.H.; Kim, S.S.
Biochim. Biophys. Acta 1157, 178-184, 1993
A;Title: Purification and characterization of the valine sensitive acetolactate synthase
A;Reference number: S33781; MUID:93283409; PMID:8507653
A;Accession: S33782
A;Molecule type: protein
A;Residues: 1-11 <YAN>
A;Experimental source: ATCC 25419
C;Complex: heterotetramer; two small and two large chains
C;Function:
A;Description: catalyzes the condensation of pyruvate and alpha-ketobutyrate to form alph
A;Pathway: valine, leucine, and isoleucine biosynthesis
A;Note: this isoenzyme exhibits homotropic allosterism with pyruvate
C;Keywords: branched-chain amino acid biosynthesis; carbon-carbon lyase; flavoprotein; he

  Query Match             27.6%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 MFLLSRN 9
              | |: ||
Db          1 MILVGRN 7


RESULT 8
PH0915
T-cell receptor beta chain V-D-J region (isolate 1) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
```

```
C;Accession: PH0915
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergi
A;Reference number: PH0891; MUID:92078857; PMID:1836012
A;Accession: PH0915
A;Molecule type: mRNA
A;Residues: 1-14 <GOL>
A;Experimental source: concanavalin A-activated lymphoblast
A;Note: the authors translated the codon GGG for residue 8 as Glu and GAG for residue 9
C;Keywords: T-cell receptor

  Query Match             27.6%;  Score 21;  DB 2;  Length 14;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          8 RNTGE 12
              | |||
Db          5 RGTGE 9


RESULT 9
A27696
contraction-inhibiting peptide I - blue mussel
C;Species: Mytilus edulis (blue mussel)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A27696
R;Hirata, T.; Kubota, I.; Iwasawa, N.; Takabatake, I.; Ikeda, T.; Muneoka, Y.
Biochem. Biophys. Res. Commun. 152, 1376-1382, 1988
A;Title: Structures and actions of Mytilus inhibitory peptides.
A;Reference number: A90142; MUID:88240357; PMID:3377776
A;Accession: A27696
A;Molecule type: protein
A;Residues: 1-6 <HIR>
A;Cross-references: UNIPROT:P13736
C;Keywords: amidated carboxyl end
F;6/Modified site: amidated carboxyl end (Val) #status experimental

  Query Match             26.3%;  Score 20;  DB 2;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPMFL 5
              :|||:
Db          2 SPMFV 6


RESULT 10
PT0691
T-cell receptor beta chain V-D-J region (154-2K) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0691
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0691
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-8 <FEE>
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             26.3%;  Score 20;  DB 2;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 TGEVRT 15
              :|| ||
Db          2 SGEPRT 7


RESULT 11
B41835
translation elongation factor EF-G homolog - Bacillus subtilis (fragment)
C;Species: Bacillus subtilis
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Dec-1997
C;Accession: B41835
R;Mitchell, C.; Morris, P.W.; Vary, J.C.
J. Bacteriol. 174, 2474-2477, 1992
A;Title: Identification of proteins phosphorylated by ATP during sporulation of Bacillus
A;Reference number: A41835; MUID:92210489; PMID:1556067
A;Accession: B41835
A;Molecule type: protein
A;Residues: 1-11 <MIT>
A;Note: this protein is phosphorylated during stationary phase but not during exponential
C;Keywords: phosphoprotein

  Query Match             26.3%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 2.1e+03;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    2;  Gaps    1;

Qy          4 FLL--SRNTG 11
              |||  :|| |
Db          1 FLLEKTRNIG 10


RESULT 12
G49410
t-complex polypeptide 1 homolog (peak 6b fraction) - rabbit (fragment)
N;Alternate names: chaperonin homolog (peak 6b)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 04-Sep-1998
C;Accession: G49410
R;Rommelaere, H.; Van Troys, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.; Amp
Proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993
A;Title: Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven relat
A;Reference number: A49410; MUID:94089752; PMID:7903455
A;Accession: G49410
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <ROM>
A;Experimental source: reticulocyte
C;Superfamily: molecular chaperone t-complex-type

  Query Match             26.3%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   57.1%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 LLSRNTG 11
              :|  |||
Db          1 ILIANTG 7


RESULT 13
PH0750
T-cell receptor beta chain (C11) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0750
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-re
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0750
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X06841
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             26.3%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          9 NTGEV 13
              |||::
Db          8 NTGQL 12
```

RESULT 14
PC4213
bphB protein - Comamonas testosteroni (fragment)
C;Species: Comamonas testosteroni
C;Date: 17-Dec-1996 #sequence_revision 21-Jan-1997 #text_change 09-Jul-2004
C;Accession: PC4213
R;Sylvestre, M.; Sirois, M.; Hurtubise, Y.; Bergeron, J.; Ahmad, D.; Shareck, F.; Barria
Gene 174, 195-202, 1996
A;Title: Sequencing of Comamonas testosteroni strain B-356-biphenyl/chlorobiphenyl dioxy
A;Reference number: JC4993; MUID:97045812; PMID:8890734
A;Accession: PC4213
A;Molecule type: DNA
A;Residues: 1-15 <SYL>
A;Cross-references: UNIPROT:Q46381; GB:U47637; NID:g1245151; PIDN:AAC44530.1; PID:g12451
A;Experimental source: strain B-356
C;Genetics:
A;Gene: bphB

```
  Query Match             26.3%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 TGEV 13
              ||||
Db          4 TGEV 7
```

RESULT 15
B27696
contraction-inhibiting peptide II - blue mussel
C;Species: Mytilus edulis (blue mussel)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: B27696
R;Hirata, T.; Kubota, I.; Iwasawa, N.; Takabatake, I.; Ikeda, T.; Muneoka, Y.
Biochem. Biophys. Res. Commun. 152, 1376-1382, 1988
A;Title: Structures and actions of Mytilus inhibitory peptides.
A;Reference number: A90142; MUID:88240357; PMID:3377776
A;Accession: B27696
A;Molecule type: protein
A;Residues: 1-6 <HIR>
A;Cross-references: UNIPROT:P13737
C;Keywords: amidated carboxyl end
F;6/Modified site: amidated carboxyl end (Val) #status experimental

```
  Query Match             25.0%;  Score 19;  DB 2;  Length 6;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 PMFL 5
              |||:
Db          3 PMFV 6
```

Search completed: February 22, 2005, 09:46:25
Job time : 12.1333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-4
Perfect score: 76
Sequence: 1 TPMFLLSRNTGEVRT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	34.2	13	2	Q9X3E1	Q9x3e1 prochloroco
2	24	31.6	10	2	Q8SHB1	Q8shb1 rhampholeon
3	23	30.3	15	2	Q7M4Z5	Q7m4z5 fusarium sp
4	23	30.3	15	2	Q7RCA8	Q7rca8 plasmodium
5	22	28.9	14	2	Q6LCR2	Q6lcr2 mus musculu
6	22	28.9	15	2	Q7M4Z4	Q7m4z4 fusarium sp
7	22	28.9	15	2	Q842L7	Q842l7 salmonella
8	21	27.6	8	2	Q659Q3	Q659q3 silene oste
9	21	27.6	8	2	Q659Q5	Q659q5 silene invo
10	21	27.6	9	1	UPA6_HUMAN	P30092 homo sapien
11	20	26.3	6	1	CIP1_MYTED	P13736 mytilus edu
12	20	26.3	8	2	Q9RQ57	Q9rq57 buchnera ap
13	20	26.3	9	2	Q9S8J8	Q9s8j8 oryza sativ
14	20	26.3	10	2	Q6UJI8	Q6uji8 nephrurus l
15	20	26.3	10	2	Q6UJJ1	Q6ujj1 nephrurus v
16	20	26.3	13	2	Q13377	Q13377 homo sapien
17	20	26.3	13	2	Q8RSG9	Q8rsg9 uncultured
18	20	26.3	13	2	Q9X3J6	Q9x3j6 prochloroco
19	20	26.3	13	2	Q7ZZI8	Q7zzi8 motacilla f
20	20	26.3	15	1	MILT_ONCKE	P81037 oncorhynchu
21	19.5	25.7	15	2	Q9TNP0	Q9tnp0 mus sp. . m
22	19	25.0	6	1	CIP2_MYTED	P13737 mytilus edu
23	19	25.0	6	1	UN06_CLOPA	P81351 clostridium
24	19	25.0	10	1	TKS1_AEDAE	P42634 aedes aegyp
25	19	25.0	10	2	Q7RKS4	Q7rks4 plasmodium
26	19	25.0	10	2	Q9TRU6	Q9tru6 bos taurus
27	19	25.0	10	2	Q8SH99	Q8sh99 brookesia n
28	19	25.0	10	2	Q8SHK1	Q8shk1 calumma nas
29	19	25.0	10	2	Q8SHL0	Q8shl0 calumma gas
30	19	25.0	10	2	Q8SHL3	Q8shl3 calumma fur
31	19	25.0	10	2	Q8SHM2	Q8shm2 calumma boe
32	19	25.0	10	2	Q8SHP0	Q8shp0 bradypodion
33	19	25.0	11	2	Q44237	Q44237 anabaena sp
34	19	25.0	12	2	Q9XNR6	Q9xnr6 pylaiella l
35	19	25.0	14	2	Q6I7N3	Q6i7n3 haemaphysal
36	19	25.0	15	1	UE15_HORVU	P34938 hordeum vul
37	18	23.7	10	2	Q7Y0R0	Q7y0r0 brassica ju
38	18	23.7	11	2	Q8IVG8	Q8ivg8 homo sapien
39	18	23.7	11	2	Q6SB14	Q6sb14 drosophila
40	18	23.7	11	2	Q997C1	Q997c1 east africa
41	18	23.7	12	2	Q8J1G6	Q8j1g6 ashbya goss
42	18	23.7	12	2	Q9P116	Q9p116 homo sapien
43	18	23.7	12	2	Q41856	Q41856 zea mays (m
44	18	23.7	12	2	Q83U71	Q83u71 salmonella
45	18	23.7	13	2	O75905	O75905 homo sapien

ALIGNMENTS

RESULT 1
Q9X3E1
ID Q9X3E1 PRELIMINARY; PRT; 13 AA.
AC Q9X3E1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Cytochrome b (Fragment).
GN Name=petB;
OS Prochlorococcus sp.
OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=1220;
RN [1]
RP SEQUENCE FROM N.A.
RA Urbach E., Chisholm S.W.;
RT "Genetic diversity in Prochlorococcus populations flow cytometrically
RT sorted from the Sargasso Sea and Gulf Stream.";
RL Limnol. Oceanog. 43:1615-1630(1998).
DR EMBL; AF070141; AAD20755.1; -.
FT NON_TER 1 1
SQ SEQUENCE 13 AA; 1434 MW; 57EBE8029A8666D3 CRC64;

Query Match 34.2%; Score 26; DB 2; Length 13;
Best Local Similarity 62.5%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 FLLSRNTG 11
              ||:||  |
Db          1 FLMSRKQG 8
```

RESULT 2
Q8SHB1
ID Q8SHB1 PRELIMINARY; PRT; 10 AA.
AC Q8SHB1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cytochrome c oxidase subunit I (Fragment).
GN Name=COI;
OS Rhampholeon brevicaudatus (Bearded pygmy chameleon).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae;
OC Rhampholeon.
OX NCBI_TaxID=91912;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22169767; PubMed=12182400; DOI=10.1006/mpev.2001.1076;
RA Townsend T., Larson A.;
RT "Molecular phylogenetics and mitochondrial genomic evolution in the
RT chamaeleonidae (Reptilia, Squamata).";
RL Mol. Phylogenet. Evol. 23:22-36(2002).

```
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Townsend T.M., Larson A.L.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF448771; AAL90595.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1291 MW;  86218E2733641771 CRC64;

  Query Match             31.6%;  Score 24;  DB 2;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 1.9e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TPMFLLS 7
              || :|||
Db          3 TPRWLLS 9


RESULT 3
Q7M4Z5
ID   Q7M4Z5      PRELIMINARY;      PRT;    15 AA.
AC   Q7M4Z5;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Fructose-bisphosphate aldolase (EC 4.1.2.13) I (Fragment).
OS   Fusarium sporotrichioides.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
OX   NCBI_TaxID=5514;
RN   [1]
RP   SEQUENCE.
RA   Chow L.P., Fukaya N., Sugiura Y., Ueno Y., Tabuchi K., Tsugita A.;
RL   Submitted (OCT-1994) to the PIR data bank.
DR   PIR; PA0075; PA0075.
DR   GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1515 MW;  74E3FE999D72632F CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 4.6e+03;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 LLSRNTGEV 13
              :|||::| :
Db          5 VLSRDSGVI 13


RESULT 4
Q7RCA8
ID   Q7RCA8      PRELIMINARY;      PRT;    15 AA.
AC   Q7RCA8;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   Name=PY05876;
OS   Plasmodium yoelii yoelii.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=73239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=17XNL;
RX   PubMed=12368865; DOI=10.1038/nature01099;
RA   Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA   Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA   Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA   Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA   Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA   Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA   Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA   van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA   Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA   Carucci D.J.;
RT   "Genome sequence and comparative analysis of the model rodent malaria
RT   parasite Plasmodium yoelii yoelii.";
RL   Nature 419:512-519(2002).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABL01001927; EAA17977.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1730 MW;  9CF62D019481ACF8 CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 4.6e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 TPMFLLS 7
              :|:| ||
Db          2 SPLFYLS 8


RESULT 5
Q6LCR2
ID   Q6LCR2      PRELIMINARY;      PRT;    14 AA.
AC   Q6LCR2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Short-chain acyl-CoA dehydrogenase (Fragment).
GN   Name=Acads;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SV;
RX   MEDLINE=96325094; PubMed=8661694;
RA   Kelly C.L., Wood P.A.;
RT   "Cloning and characterization of the mouse short-chain acyl-CoA
RT   dehydrogenase gene.";
RL   Mamm. Genome 7:262-264(1996).
DR   EMBL; U36274; AAB99744.1; -.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1467 MW;  DF251B3D3157FD0D CRC64;

  Query Match             28.9%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 6.4e+03;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 LLSRNTGEVR 14
              ||:|  | :|
Db          5 LLARARGPLR 14


RESULT 6
Q7M4Z4
ID   Q7M4Z4      PRELIMINARY;      PRT;    15 AA.
AC   Q7M4Z4;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Fructose-bisphosphate aldolase (EC 4.1.2.13) III (Fragment).
OS   Fusarium sporotrichioides.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
OX   NCBI_TaxID=5514;
RN   [1]
RP   SEQUENCE.
RA   Chow L.P., Fukaya N., Sugiura Y., Ueno Y., Tabuchi K., Tsugita A.;
RL   Submitted (OCT-1994) to the PIR data bank.
```

```
DR   PIR; PA0102; PA0102.
DR   GO; GO:0004332; F:fructose-bisphosphate aldolase activity; IEA.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1528 MW;  74E3FE999D726ABF CRC64;

  Query Match             28.9%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 6.9e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 LLSRNTGEV 13
              :||| :| :
Db          5 VLSRKSGVI 13


RESULT 7
Q842L7
ID   Q842L7      PRELIMINARY;      PRT;    15 AA.
AC   Q842L7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   QacEdelta1 (Fragment).
GN   Name=qacEdelta1;
OS   Salmonella enterica subsp. enterica serovar Typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=90370;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YMC95/6/4405;
RA   Yum J.H., Yong D., Lee K., Chong Y.;
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY245101; AAO89217.1; -.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1675 MW;  C2E1985D5C9E287B CRC64;

  Query Match             28.9%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 6.9e+03;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 MFLLSRNTGEV 13
              :||:    |||
Db          5 LFLVIAIVGEV 15


RESULT 8
Q659Q3
ID   Q659Q3      PRELIMINARY;      PRT;     8 AA.
AC   Q659Q3;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Photo system II reaction centre subunit V (Fragment).
GN   Name=psbE;
OS   Silene ostenfeldii.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Caryophyllaceae; Silene.
OX   NCBI_TaxID=269081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Popp M., Erixon P., Eggens F., Oxelman B.;
RT   "Origin and Evolution of a Circumpolar Polyploid Species Complex in
RT   Silene (Caryophyllaceae) Inferred from Low Copy Nuclear RNA Polymerase
RT   Introns, rDNA, and Chloroplast DNA.";
RL   Syst. Bot. 0:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Popp M.;
RT   "Disentangling the reticulate history of polyploids in Silene
RT   (Caryophyllaceae).";
RL   Thesis (2004), Department of Evolutionary Biology, Systematic Botany,
RL   Uppsala University, Uppsala, Sweden.
DR   EMBL; AJ831757; CAH41967.1; -.
KW   Chloroplast.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  824 MW;  ED0B1861B5B865B6 CRC64;

  Query Match             27.6%;  Score 21;  DB 2;  Length 8;
  Best Local Similarity   57.1%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LSRNTGE 12
              :| :|||
Db          1 MSGSTGE 7


RESULT 9
Q659Q5
ID   Q659Q5      PRELIMINARY;      PRT;     8 AA.
AC   Q659Q5;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Photo system II reaction centre subunit V (Fragment).
GN   Name=psbE;
OS   Silene involucrata.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Caryophyllaceae; Silene.
OX   NCBI_TaxID=39884;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Popp M., Erixon P., Eggens F., Oxelman B.;
RT   "Origin and Evolution of a Circumpolar Polyploid Species Complex in
RT   Silene (Caryophyllaceae) Inferred from Low Copy Nuclear RNA Polymerase
RT   Introns, rDNA, and Chloroplast DNA.";
RL   Syst. Bot. 0:0-0(0).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RA   Popp M.;
RT   "Disentangling the reticulate history of polyploids in Silene
RT   (Caryophyllaceae).";
RL   Thesis (2004), Department of Evolutionary Biology, Systematic Botany,
RL   Uppsala University, Uppsala, Sweden.
DR   EMBL; AJ831755; CAH41963.1; -.
KW   Chloroplast.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  824 MW;  ED0B1861B5B865B6 CRC64;

  Query Match             27.6%;  Score 21;  DB 2;  Length 8;
  Best Local Similarity   57.1%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 LSRNTGE 12
              :| :|||
Db          1 MSGSTGE 7


RESULT 10
UPA6_HUMAN
ID   UPA6_HUMAN     STANDARD;      PRT;     9 AA.
AC   P30092;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Unknown protein from 2D-PAGE of plasma (Spot 14) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=93092937; PubMed=1459097;
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 5, its MW is: 48 kDa.
DR   SWISS-2DPAGE; P30092; HUMAN.
KW   Direct protein sequencing.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  935 MW;  5282F2CAA8676447 CRC64;

  Query Match             27.6%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 NTGEVRT 15
              | |:| |
Db          3 NPGDVXT 9


RESULT 11
CIP1_MYTED
ID   CIP1_MYTED     STANDARD;      PRT;     6 AA.
AC   P13736;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Contraction-inhibiting peptide I (MIP I).
OS   Mytilus edulis (Blue mussel).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;
OC   Mytiloidea; Mytilidae; Mytilus.
OX   NCBI_TaxID=6550;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Pedal ganglion;
RX   MEDLINE=88240357; PubMed=3377776;
RA   Hirata T., Kubota I., Iwasawa N., Takabatake I., Ikeda T., Muneoka Y.;
RT   "Structures and actions of Mytilus inhibitory peptides.";
RL   Biochem. Biophys. Res. Commun. 152:1376-1382(1988).
CC   -!- FUNCTION: Inhibitory action on contractions in several molluscan
CC       muscles.
CC   -!- SIMILARITY: To MIP II.
DR   PIR; A27696; A27696.
KW   Amidation; Direct protein sequencing; Hormone.
FT   MOD_RES       6      6       Valine amide.
SQ   SEQUENCE   6 AA;  637 MW;  72C9C68775B81000 CRC64;

  Query Match             26.3%;  Score 20;  DB 1;  Length 6;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPMFL 5
              :|||:
Db          2 SPMFV 6


RESULT 12
Q9RQ57
ID   Q9RQ57      PRELIMINARY;      PRT;     8 AA.
AC   Q9RQ57;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   NifS protein homolog (Fragment).
GN   Name=nifS;
OS   Buchnera aphidicola.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Buchnera.
OX   NCBI_TaxID=9;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20022990; PubMed=10555290;
RA   Clark M.A., Moran N.A., Baumann P.;
RT   "Sequence evolution in bacterial endosymbionts having extreme base
RT   compositions.";
RL   Mol. Biol. Evol. 16:1586-1598(1999).
DR   EMBL; AF130812; AAF13797.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  980 MW;  F3A73B504771A336 CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPMFL 5
              ||::|
Db          3 TPIYL 7


RESULT 13
Q9S8J8
ID   Q9S8J8      PRELIMINARY;      PRT;     9 AA.
AC   Q9S8J8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   ORYZATENSIN=BIOACTIVE peptide.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95102521; PubMed=7804141;
RA   Takahashi M., Moriguchi S., Yoshikawa M., Sasaki R.;
RT   "Isolation and characterization of oryzatensin: a novel bioactive
RT   peptide with ileum-contracting and immunomodulating activities derived
RT   from rice albumin.";
RL   Biochem. Mol. Biol. Int. 33:1151-1158(1994).
DR   Gramene; Q9S8J8; -.
SQ   SEQUENCE   9 AA;  1093 MW;  0E8C67377B56877B CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 PMFLLSR 8
              ||: | |
Db          3 PMYPLPR 9


RESULT 14
Q6UJI8
ID   Q6UJI8      PRELIMINARY;      PRT;    10 AA.
AC   Q6UJI8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Nephrurus laevissimis.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae; Nephrurus.
OX   NCBI_TaxID=255191;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Melville J., Schulte J.A. II, Larson A.;
RT   "A Molecular Study of Phylogenetic Relationships and Evolution of
RT   Antipredator Strategies in Australian Diplodactylus Geckos, Subgenus
RT   Strophurus.";
RL   Biol. J. Linn. Soc. Lond. 82:123-138(2004).
DR   EMBL; AY369020; AAR18892.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1270 MW;  9AD380C733641771 CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   57.1%;  Pred. No. 1e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TPMFLLS 7
              || :| |
Db          3 TPRWLFS 9


RESULT 15
Q6UJJ1
ID   Q6UJJ1      PRELIMINARY;      PRT;    10 AA.
AC   Q6UJJ1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Nephrurus vertebralis.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Gekkota; Gekkonidae; Nephrurus.
OX   NCBI_TaxID=255190;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Melville J., Schulte J.A. II, Larson A.;
RT   "A Molecular Study of Phylogenetic Relationships and Evolution of
RT   Antipredator Strategies in Australian Diplodactylus Geckos, Subgenus
RT   Strophurus.";
RL   Biol. J. Linn. Soc. Lond. 82:123-138(2004).
DR   EMBL; AY369019; AAR18889.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1270 MW;  9AD380C733641771 CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   57.1%;  Pred. No. 1e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TPMFLLS 7
              || :| |
Db          3 TPRWLFS 9


Search completed: February 22, 2005, 09:37:51
Job time : 54.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-5
Perfect score: 88
Sequence: 1 CECNIKVKDVNDNFP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	31.8	12	2	S71034	potB protein - Sal
2	24	27.3	15	2	PC2215	fibrinogenolytic p
3	23.5	26.7	14	2	S41601	interferon alpha r
4	23	26.1	5	2	A33882	cadmium-binding pe
5	23	26.1	7	2	B33882	cadmium-binding he
6	23	26.1	14	2	A28018	very late antigen-
7	22	25.0	13	2	S47362	T-cell antigen rec
8	22	25.0	13	2	A61210	antibiotic GE2270
9	21	23.9	11	2	S54347	tubulin beta chain
10	21	23.9	15	2	A47146	topoisomerase I -
11	21	23.9	15	2	PS0212	29K protein 4228 -
12	20	22.7	15	2	S57584	T cell receptor V-
13	19	21.6	8	2	PL0184	capsid protein VP-
14	19	21.6	9	2	A61230	calsequestrin, car
15	19	21.6	9	2	S66635	alpha-2-macroglobu
16	19	21.6	10	2	S51912	hemagglutinin - in
17	19	21.6	11	2	PH1376	T antigen variant
18	19	21.6	12	2	PN0160	ribosomal protein
19	19	21.6	15	2	A54397	ubiquitin-carrier
20	19	21.6	15	2	D48394	major fat-globule
21	19	21.6	15	2	C61511	milk band B protei
22	19	21.6	15	2	G49255	T-cell receptor be
23	18	20.5	7	2	B44787	calliFMRFamide 11
24	18	20.5	8	2	S59622	metallothionein is
25	18	20.5	10	2	PH0933	T-cell receptor be
26	18	20.5	10	2	C44787	calliFMRFamide 12
27	18	20.5	10	2	B49581	sialokinin II - ye
28	18	20.5	11	2	T06383	hypothetical prote
29	18	20.5	11	2	PH1375	T antigen variant
30	18	20.5	12	2	A49033	T-cell receptor de
31	18	20.5	12	2	B49033	T-cell receptor de
32	18	20.5	12	2	JQ2308	hypothetical 1.4K
33	18	20.5	12	2	JQ2318	hypothetical 1.4K
34	18	20.5	13	2	S47392	T-cell antigen rec
35	18	20.5	13	2	B25448	Ig kappa-1 chain,
36	18	20.5	13	2	JH0460	corticostatic pept
37	18	20.5	14	2	S29486	GTP-binding protei
38	18	20.5	15	2	A49252	T-cell receptor be
39	18	20.5	15	2	PN0118	hemoglobin beta ch
40	17.5	19.9	13	2	S29488	GTP-binding protei
41	17	19.3	8	2	S19288	acylase - Kluyvera
42	17	19.3	10	1	SPPGNK	neuromedin K - pig
43	17	19.3	10	2	S27873	hypothetical prote
44	17	19.3	11	2	B49164	chromogranin-B - r
45	17	19.3	11	2	D58502	27K bile and gallb

ALIGNMENTS

RESULT 1
S71034
potB protein - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S71034
R;Stein, M.A.; Leung, K.Y.; Zwick, M.; Garcia-del Portillo, F.; Finlay, B.B.
Mol. Microbiol. 20, 151-164, 1996
A;Title: Identification of a Salmonella virulence gene required for formation of filament
A;Reference number: S71033; MUID:97014378; PMID:8861213
A;Accession: S71034
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-12 <STE>
A;Cross-references: UNIPROT:Q56060; EMBL:U51867; NID:g1272352; PIDN:AAA97466.1; PID:g127
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1996
C;Genetics:
A;Gene: potB

Query Match 31.8%; Score 28; DB 2; Length 12;
Best Local Similarity 55.6%; Pred. No. 2.7e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 NIKVKDVND 12
              | || |::|
Db          4 NKKVSDISD 12
```

RESULT 2
PC2215
fibrinogenolytic proteinase A2 (EC 3.4.21.-) - western diamondback rattlesnake (fragment)
N;Alternate names: alpha-fibrinogenase A2
C;Species: Crotalus atrox (western diamondback rattlesnake)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PC2215
R;Hung, C.C.; Chiou, S.H.
Biochem. Biophys. Res. Commun. 201, 1414-1423, 1994
A;Title: Isolation of multiple isoforms of alpha-fibrinogenase from the western diamondb
viper.
A;Reference number: PC2214; MUID:94296418; PMID:8024586
A;Accession: PC2215
A;Molecule type: protein
A;Residues: 1-15 <HUN>
A;Cross-references: UNIPROT:Q9PRW3
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase

Query Match 27.3%; Score 24; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 ECNI 5
```

```
              ||||
Db          6 ECNI 9


RESULT 3
S41601
interferon alpha receptor 1 - human (fragments)
C;Species: Homo sapiens (man)
C;Date: 25-Dec-1994 #sequence_revision 01-Dec-1995 #text_change 30-May-1997
C;Accession: S41601
R;Abramovich, C.; Ratovitski, E.; Lundgren, E.; Revel, M.
FEBS Lett. 338, 295-300, 1994
A;Title: Identification of mRNAs encoding two different soluble forms of the human inter
A;Reference number: S41601; MUID:94139943; PMID:8307198
A;Accession: S41601
A;Molecule type: mRNA
A;Residues: 1-14 <ABR>
C;Keywords: cytokine receptor

  Query Match             26.7%;  Score 23.5;  DB 2;  Length 14;
  Best Local Similarity   62.5%;  Pred. No. 1.6e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          1 CEC-NIKV 7
              ||| || :
Db          4 CECENISL 11


RESULT 4
A33882
cadmium-binding pentapeptide - downy thornapple
C;Species: Datura innoxia (downy thornapple)
C;Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 18-Jun-1993
C;Accession: A33882
R;Jackson, P.J.; Unkefer, C.J.; Doolen, J.A.; Watt, K.; Robinson, N.J.
Proc. Natl. Acad. Sci. U.S.A. 84, 6619-6623, 1987
A;Title: Poly(gamma-glutamylcysteinyl)glycine: its role in cadmium resistance in plant c
A;Reference number: A94182; MUID:88016144; PMID:3477793
A;Accession: A33882
A;Molecule type: protein
A;Residues: 1-5 <JAC>

  Query Match             26.1%;  Score 23;  DB 2;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CEC 3
              |||
Db          2 CEC 4


RESULT 5
B33882
cadmium-binding heptapeptide - downy thornapple
C;Species: Datura innoxia (downy thornapple)
C;Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 18-Jun-1993
C;Accession: B33882
R;Jackson, P.J.; Unkefer, C.J.; Doolen, J.A.; Watt, K.; Robinson, N.J.
Proc. Natl. Acad. Sci. U.S.A. 84, 6619-6623, 1987
A;Title: Poly(gamma-glutamylcysteinyl)glycine: its role in cadmium resistance in plant c
A;Reference number: A94182; MUID:88016144; PMID:3477793
A;Accession: B33882
A;Molecule type: protein
A;Residues: 1-7 <JA2>

  Query Match             26.1%;  Score 23;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CEC 3
              |||
Db          2 CEC 4


RESULT 6
A28018
very late antigen-1 alpha chain - human (fragment)
N;Alternate names: VLA-1 alpha chain
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 12-May-1994
C;Accession: A28018
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987
A;Title: The very late antigen family of heterodimers is part of a superfamily of molecul
A;Reference number: A94151; MUID:87204112; PMID:3033641
A;Accession: A28018
A;Molecule type: protein
A;Residues: 1-14 <TAK>
C;Keywords: duplication; heterodimer; membrane protein

  Query Match             26.1%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 NIKVKD 9
              |: |||
Db          2 NVDVKD 7


RESULT 7
S47362
T-cell antigen receptor VJ junction beta chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: S47362
R;Lehner, P.J.
submitted to the EMBL Data Library, August 1994
A;Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T ce
A;Reference number: S47355
A;Accession: S47362
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LEH>
A;Cross-references: EMBL:Z35688; NID:g527465; PIDN:CAA84757.1; PID:g527466
C;Keywords: T-cell receptor

  Query Match             25.0%;  Score 22;  DB 2;  Length 13;
  Best Local Similarity   33.3%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 CNIKVKDVNDNF 14
              |   |   |: |
Db          1 CASSVSSYNEQF 12


RESULT 8
A61210
antibiotic GE2270 - Planobispora rosea (ATCC 53773)
C;Species: Planobispora rosea
C;Date: 13-May-1994 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C;Accession: A61210
R;Kettenring, J.; Colombo, L.; Ferrari, P.; Tavecchia, P.; Nebuloni, M.; Vekey, K.; Gallo
J. Antibiot. 44, 702-715, 1991
A;Title: Antibiotic GE2270 A: a novel inhibitor of bacterial protein synthesis. II. Struc
A;Reference number: A61210; MUID:91349090; PMID:1880060
A;Accession: A61210
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-13 <KET>
A;Cross-references: UNIPROT:Q7M0J8
C;Keywords: amidated carboxyl end; methylated amino acid; oxazole/thiazole ring
F;1-10/Cross-link: thiazole amino end (Cys-Lys) (by 10-C6) #status experimental
F;2-3/Cross-link: 5-methoxythiazole (Val-Cys) #status experimental
F;4-5/Cross-link: 5-methylthiazole (Asn-Cys) #status experimental
F;4/Modified site: N4-methylasparagine (Asn) #status experimental
F;7-8/Cross-link: thiazole (Phe-Cys) #status experimental
```

```
F;7/Modified site: 3-hydroxyphenylalanine (Phe) #status experimental
F;8-9/Cross-link: thiazole (Cys-Cys) #status experimental
F;9-10/Cross-link: pyridine (Cys-Lys) #status experimental
F;10-11/Cross-link: thiazole (Lys-Cys) #status experimental
F;11-12/Cross-link: 2-oxazoline (Cys-Ser) #status experimental
F;13/Modified site: amidated carboxyl end (Pro) #status experimental

  Query Match             25.0%;  Score 22;  DB 2;  Length 13;
  Best Local Similarity   75.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CECN 4
              | ||
Db          1 CVCN 4


RESULT 9
S54347
tubulin beta chain - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 27-Oct-1995 #sequence_revision 30-Jan-1998 #text_change 09-Jul-2004
C;Accession: S54347
R;Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.
Biochem. J. 306, 551-555, 1995
A;Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia
A;Reference number: S54343; MUID:95194333; PMID:7887910
A;Accession: S54347
A;Molecule type: protein
A;Residues: 1-11 <OKA>
A;Cross-references: UNIPROT:Q7M372

  Query Match             23.9%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 KDVN 11
              ||||
Db          1 KDVN 4


RESULT 10
A47146
topoisomerase I - vaccinia virus (fragment)
C;Species: vaccinia virus
C;Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 09-Jul-2004
C;Accession: A47146
R;Klemperer, N.; Traktman, P.
J. Biol. Chem. 268, 15887-15899, 1993
A;Title: Biochemical analysis of mutant alleles of the vaccinia virus topoisomerase I ca
A;Reference number: A47146; MUID:93340198; PMID:8393454
A;Accession: A47146
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-15 <KLE>
A;Cross-references: UNIPROT:Q9JFB0; GB:L13447
C;Superfamily: vaccinia virus DNA topoisomerase

  Query Match             23.9%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 4.3e+03;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 IKVKDV 10
              |::||:
Db          2 IRIKDL 7


RESULT 11
PS0212
29K protein 4228 - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: PS0212
R;Tsugita, A.; Miyatake, N.
submitted to JIPID, April 1993
A;Reference number: PS0208
A;Accession: PS0212
A;Molecule type: protein
A;Residues: 1-15 <TSU>
A;Cross-references: UNIPROT:Q7M279
A;Experimental source: germ
C;Comment: molecular weight 29K, pI 6.1.

  Query Match             23.9%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 4.3e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 KDVNDNF 14
              :|| | :
Db          7 RDVGDRY 13


RESULT 12
S57584
T cell receptor V-D-J junctional alpha chain region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57584
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified by
A;Reference number: S57494
A;Accession: S57584
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <BUR>
A;Cross-references: EMBL:Z49956; NID:g887466; PIDN:CAA90227.1; PID:g887467
C;Keywords: T-cell receptor

  Query Match             22.7%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   33.3%;  Pred. No. 6.2e+03;
  Matches    3;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CECNIKVKD 9
              |  | : :|
Db          1 CAVNARERD 9


RESULT 13
PL0184
capsid protein VP-1 - murine poliovirus (fragment)
C;Species: murine poliovirus, Theiler's encephalomyelitis virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
C;Accession: PL0184
R;Zurbriggen, A.; Hogle, J.M.; Fujinami, R.S.
J. Exp. Med. 170, 2037-2049, 1989
A;Title: Alteration of amino acid 101 within capsid protein VP-1 changes the pathogenicit
A;Reference number: PL0184; MUID:90063468; PMID:2479706
A;Accession: PL0184
A;Molecule type: genomic RNA
A;Residues: 1-8 <ZUR>
C;Keywords: capsid protein

  Query Match             21.6%;  Score 19;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 NFP 15
              |||
Db          6 NFP 8


RESULT 14
A61230
calsequestrin, cardiac and slow skeletal muscle - northern leopard frog (fragment)
N;Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; laminin-b
C;Species: Rana pipiens (northern leopard frog)
```

```
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 09-Jul-2004
C;Accession: A61230
R;McLeod, A.G.; Shen, A.C.Y.; Campbell, K.P.; Michalak, M.; Jorgensen, A.O.
Circ. Res. 69, 344-359, 1991
A;Title: Frog cardiac calsequestrin. Identification, characterization, and subcellular d
rdium.
A;Reference number: A61230; MUID:91316784; PMID:1860177
A;Accession: A61230
A;Molecule type: protein
A;Residues: 1-9 <MCL>
A;Cross-references: UNIPROT:Q7LZ81
C;Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding protei
C;Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound to c
C;Comment: The fast skeletal muscle isoform of calsequestrin can be phosphorylated in vi
C;Superfamily: calsequestrin
C;Keywords: calcium binding; cardiac muscle; glycoprotein; heart; phosphoprotein; skelet

  Query Match             21.6%;  Score 19;  DB 2;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         13 NFP 15
              |||
Db          5 NFP 7


RESULT 15
S66635
alpha-2-macroglobulin isoform 1 - bovine (fragment)
C;Species: Bos primigenius indicus (zebu cattle)
C;Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66635
R;Dolmer, K.; Jenner, L.B.; Jacobsen, L.; Andersen, G.R.; Koch, T.J.; Thirup, S.; Sottru
FEBS Lett. 372, 93-95, 1995
A;Title: Crystallisation and preliminary X-ray analysis of the receptor-binding domain c
A;Reference number: S66634; MUID:96032553; PMID:7556651
A;Accession: S66635
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <DOL>
A;Cross-references: UNIPROT:Q7M2N8

  Query Match             21.6%;  Score 19;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         12 DNFP 15
              | ||
Db          2 DEFP 5


Search completed: February 22, 2005, 09:46:25
Job time : 11.1333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-4
Perfect score: 76
Sequence: 1 TPMFLLSRNTGEVRT 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	76	100.0	15	2	AAW04844	Aaw04844 Self epit
2	76	100.0	15	2	AAW64816	Aaw64816 Desmoglei
3	76	100.0	15	2	AAW78815	Aaw78815 Desmoglei
4	76	100.0	15	3	AAB33626	Aab33626 MHC class
5	76	100.0	15	4	AAG93722	Aag93722 Human des
6	76	100.0	15	5	AAO17034	Aao17034 Desmoglei
7	76	100.0	15	6	ABU96578	Abu96578 MHC class
8	76	100.0	15	8	ADQ14319	Adq14319 Human des
9	76	100.0	15	8	ADR41701	Adr41701 Desmoglei
10	76	100.0	15	8	ADS14311	Ads14311 Desmoglei
11	43	56.6	11	3	AAY63823	Aay63823 Desmoglei
12	43	56.6	11	5	ABB45875	Abb45875 Desmoglei
13	40	52.6	11	3	AAY63733	Aay63733 Desmoglei
14	40	52.6	11	5	ABB45785	Abb45785 Desmoglei
15	39	51.3	11	3	AAY63778	Aay63778 Desmoglei
16	39	51.3	11	3	AAY63688	Aay63688 Desmoglei
17	39	51.3	11	5	ABB45830	Abb45830 Desmoglei
18	39	51.3	11	5	ABB45740	Abb45740 Desmoglei
19	38	50.0	9	3	AAY63862	Aay63862 Desmoglei
20	38	50.0	9	3	AAY63599	Aay63599 Desmoglei
21	38	50.0	9	5	ABB45654	Abb45654 Non-class
22	38	50.0	9	5	ABB45371	Abb45371 Desmoglei
23	38	50.0	9	5	ABB45387	Abb45387 Desmoglei
24	38	50.0	9	5	ABB45914	Abb45914 Desmoglei
25	38	50.0	10	3	AAY63820	Aay63820 Desmoglei
26	38	50.0	10	5	ABB45872	Abb45872 Desmoglei
27	38	50.0	11	3	AAY63643	Aay63643 Desmoglei
28	38	50.0	11	5	ABB45695	Abb45695 Desmoglei
29	35	46.1	10	3	AAY63730	Aay63730 Desmoglei
30	35	46.1	10	5	ABB45782	Abb45782 Desmoglei
31	35	46.1	11	3	AAY61096	Aay61096 Cadherin-
32	34	44.7	10	3	AAY63685	Aay63685 Desmoglei
33	34	44.7	10	3	AAY63775	Aay63775 Desmoglei
34	34	44.7	10	5	ABB45827	Abb45827 Desmoglei
35	34	44.7	10	5	ABB45737	Abb45737 Desmoglei
36	33	43.4	8	3	AAY63596	Aay63596 Desmoglei
37	33	43.4	8	3	AAY63859	Aay63859 Desmoglei
38	33	43.4	8	5	ABB45368	Abb45368 Desmoglei
39	33	43.4	8	5	ABB45911	Abb45911 Desmoglei
40	33	43.4	10	3	AAY61093	Aay61093 Cadherin-
41	33	43.4	10	3	AAY63640	Aay63640 Desmoglei
42	33	43.4	10	5	ABB45692	Abb45692 Desmoglei
43	33	43.4	11	3	AAY61477	Aay61477 Cadherin-
44	33	43.4	11	3	AAY61636	Aay61636 Cadherin-
45	33	43.4	15	2	AAW04842	Aaw04842 Self epit

ALIGNMENTS

RESULT 1
AAW04844
ID AAW04844 standard; peptide; 15 AA.
XX
AC AAW04844;
XX
DT 18-FEB-1997 (first entry)
XX
DE Self epitope of desmoglein 3, implicated in autoimmune disease.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Homo sapiens.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 1; Page 40; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is derived from the human desmoglein

```
CC   3 protein (amino acids 206-220) and is implicated as a self epitope in
CC   pemphigus vulgaris. Peptides derived from the human desmoglein protein
CC   are described in AAW04841-47
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 76;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPMFLLSRNTGEVRT 15
              |||||||||||||||
Db          1 TPMFLLSRNTGEVRT 15


RESULT 2
AAW64816
ID   AAW64816 standard; peptide; 15 AA.
XX
AC   AAW64816;
XX
DT   29-SEP-1998  (first entry)
XX
DE   Desmoglein-3 206-220.
XX
KW   Desmoglein; DG; gene therapy; pemphigus vulgaris; microparticle;
KW   autoantigen; autoimmune disease; MHC.
XX
OS   Homo sapiens.
XX
PN   US5783567-A.
XX
PD   21-JUL-1998.
XX
PF   22-JAN-1997;   97US-00787547.
XX
PR   22-JAN-1997;   97US-00787547.
XX
PA   (PANG-) PANGAEA PHARM INC.
XX
PI   Langer RS,  Hedley ML,  Curley JM;
XX
DR   WPI; 1998-427077/36.
XX
PT   Microparticle encapsulated nucleic acids - for recombinant expression of
PT   proteins e.g. in gene therapy.
XX
PS   Disclosure; Col 4; 42pp; English.
XX
CC   The patent describes a new preparation of microparticles each comprising
CC   a polymeric matrix and a nucleic acid. The polymeric matrix consists of
CC   one or more synthetic polymers having a solubility in water of less than
CC   1 mg/l (e.g. poly-lactic-co-glycolic acid); and at least 90% of the
CC   microparticles have a diameter of less than 100 microns. The
CC   microparticles are useful for the delivery of nucleic acids to phagocytic
CC   cells. In one embodiment the microparticles are less than 20 microns in
CC   diameter and the nucleic acid (preferably in closed circular form)
CC   includes an expression control sequence operatively linked to a coding
CC   sequence, where the expression product of the coding sequence is a
CC   polypeptide having a length and a sequence which permits it to bind to an
CC   MHC class I or II molecule. The expression product is thus an effective
CC   stimulator of an immune response in mammals. The present sequence, an
CC   antigenic portion of desmoglein 3, is an example of an MHC class II
CC   peptide which can be expressed by the nucleic acid. It is associated with
CC   pemphigus vulgaris
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 76;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPMFLLSRNTGEVRT 15
              |||||||||||||||
Db          1 TPMFLLSRNTGEVRT 15


RESULT 3
AAW78815
ID   AAW78815 standard; peptide; 15 AA.
XX
AC   AAW78815;
XX
DT   17-NOV-1998  (first entry)
XX
DE   Desmoglein 3 protein fragment 206-220.
XX
KW   Microparticle; delivery; polymeric matrix; autoantigen; tumour antigen;
KW   class II associated peptide; pathogen; gene therapy; genetic disease;
KW   infection; downregulation; immune response.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO9831398-A1.
XX
PD   23-JUL-1998.
XX
PF   22-JAN-1998;   98WO-US001499.
XX
PR   22-JAN-1997;   97US-00787547.
PR   06-JAN-1998;   98US-00003253.
XX
PA   (PANG-) PANGAEA PHARM INC.
XX
PI   Hedley ML,  Curley JM,  Langer RS,  Lunsford LB;
XX
DR   WPI; 1998-427556/36.
XX
PT   New preparations of microparticles - comprising a synthetic polymer
PT   matrix and nucleic acid comprising an expression vector for use in gene
PT   therapy.
XX
PS   Disclosure; Page 8; 101pp; English.
XX
CC   A microparticle preparation (MP) has been developed, consisting of
CC   microparticles having a diameter of less than 100 mu m. The MP comprises:
CC   (a) a polymeric matrix (PM) consisting of one or more synthetic polymers
CC   having a solubility in water of less that 1 mg/l; and (b) an expression
CC   vector selected from RNA molecules (at least 50% of which are closed
CC   circles) or circular plasmid DNA (at least 50% of which are supercoiled).
CC   Also described is a MP of at most 20 microns in diameter, comprising: (a)
CC   a PM; and (b) a NAM comprising an expression control sequence operatively
CC   linked to a coding sequence, where the coding sequence encodes an
CC   expression product selected from: (1) a polypeptide at least 7 amino
CC   acids in length, having a sequence identical to the sequence of: (i) a
CC   fragment of a naturally-occurring mammalian protein; or (ii) a fragment
CC   of a naturally-occurring protein from an infectious agent which infects a
CC   mammal; (2) a peptide having a length and sequence which permits it to
CC   bind to an MHC class I or II molecule; and (3) the polypeptide or the
CC   peptide linked to a trafficking sequence. AAW69763 to AAW69765, and
CC   AAW78793 to AAW78897 are peptide fragments for use in the present
CC   invention. The MPs are highly effective vehicles for the delivery of
CC   polynucleotides into phagocytic cells. They can be used for gene therapy,
CC   e.g. for treating genetic diseases, infections or tumours or for
CC   downregulating an immune response
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 76;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPMFLLSRNTGEVRT 15
              |||||||||||||||
Db          1 TPMFLLSRNTGEVRT 15
```

RESULT 4
AAB33626
ID AAB33626 standard; peptide; 15 AA.
XX
AC AAB33626;
XX
DT 26-JAN-2001 (first entry)
XX
DE MHC class II associated immunogenic peptide SEQ ID 25.
XX
KW Microparticle; nucleic acid delivery; immunogenic peptide; MHC I; MHC II;
KW major histocompatibility complex; vaginal tissue; mucosal tissue.
XX
OS Unidentified.
XX
PN WO200053161-A2.
XX
PD 14-SEP-2000.
XX
PF 10-MAR-2000; 2000WO-US006578.
XX
PR 11-MAR-1999; 99US-00266463.
PR 27-MAY-1999; 99US-00321346.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Lunsford LB, Putnam D, Hedley ML;
XX
DR WPI; 2000-638130/61.
XX
PT Microparticles useful for administering a nucleic acid into the mucosal
PT tissue preferably vaginal tissue of an animal, comprises a polymeric
PT matrix, a lipid and a nucleic acid molecule.
XX
PS Claim 25; Page 11; 96pp; English.
XX
CC The present invention relates to microparticles which are less than 20
CC microns in diameter, which comprise a polymeric matrix, a lipid and a
CC nucleic acid molecule. The microparticle is specifically not encapsulated
CC in a liposome and does not comprise a cell. The nucleotide sequence
CC encodes an expression product that binds to major histocompatibility
CC complex (MHC) type I or II molecules. Peptides AAB33602-B33647 represent
CC MHC class II associated immunogenic peptides, and AAB33648-B33710
CC represent MHC class I associated immunogenic peptides. The peptides are
CC examples of the expression products of the nucleotide sequences which can
CC be included in the microparticles of the invention. Sequences AAB33711-
CC B33716 represent alternative expression products and nuclear localisation
CC signals also used in the invention. The microparticles are useful for
CC administering a nucleic acid into the mucosal tissue preferably vaginal
CC tissue of an animal
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPMFLLSRNTGEVRT 15
||||||||||||||||
Db 1 TPMFLLSRNTGEVRT 15

RESULT 5
AAG93722
ID AAG93722 standard; peptide; 15 AA.
XX
AC AAG93722;
XX
DT 17-SEP-2001 (first entry)
XX
DE Human desmoglein 3 peptide 2.
XX
KW Continuous flow production; microparticle; gene therapy;
KW antisense therapy; vaccination; treatment; autoimmune disease;
KW immune response modulation.
XX
OS Homo sapiens.
XX
PN WO200136583-A1.
XX
PD 25-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-US031770.
XX
PR 19-NOV-1999; 99US-00443654.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Hsu Y, Tyo M;
XX
DR WPI; 2001-425203/45.
XX
PT Continuous production of microparticles containing nucleic acid for e.g.
PT gene therapy, comprises mixing a solution of polymeric material and
PT nucleic acid with a surfactant solution, removing solvent and drying.
XX
PS Disclosure; Page 9; 47pp; English.
XX
CC The present sequence is that of a peptide of the invention. The invention
CC relates to a method for scalable, continuous flow production of a nucleic
CC acid containing microparticle that maintains the structural integrity of
CC the associated nucleic acid and results in a microparticle having purity
CC suitable for introduction into an animal host. Microparticles prepared
CC according to the method can be used for delivery of a nucleic acid for
CC gene therapy, antisense therapy, vaccination, treatment of autoimmune
CC disease and either specific or non-specific modulation of an immune
CC response. The microparticles may also be used to deliver nucleic acid
CC encoding a protein or peptide useful in any kind of therapy. The method
CC is economical, aseptic and scalable. The method also enables control over
CC the size of microparticles. The microparticles produced are free of
CC impurities such as organic solvents and are readily dispersed in a wide
CC range of dispersing agents
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPMFLLSRNTGEVRT 15
||||||||||||||||
Db 1 TPMFLLSRNTGEVRT 15

RESULT 6
AAO17034
ID AAO17034 standard; peptide; 15 AA.
XX
AC AAO17034;
XX
DT 29-MAY-2002 (first entry)
XX
DE Desmoglein 3 residues 206-220 SEQ ID NO: 31.
XX
KW Alpha-MSH; inflammation; autoimmune disease; gene therapy; sepsis;
KW alpha-melanocyte stimulating hormone; rheumatoid arthritis; asthma;
KW cirrhosis; dermatitis; psoriasis; inflammatory bowel disease;
KW immunosuppressive; antiinflammatory; antirheumatic; antiarthritic;
KW antiasthmatic; antibacterial; dermatological; antipsoriatic;
KW antidiabetic; ophthalmological; neuroprotective; multiple sclerosis;
KW diabetes; uveitis; coeliac disease.
XX
OS Unidentified.
XX

PN WO200206316-A2.
XX
PD 24-JAN-2002.
XX
PF 16-JUL-2001; 2001WO-US022263.
XX
PR 14-JUL-2000; 2000US-0218381P.
PR 18-AUG-2000; 2000US-0226382P.
PR 06-OCT-2000; 2000US-0238380P.
PR 29-DEC-2000; 2000US-0258764P.
PR 14-JUN-2001; 2001US-0298317P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Hedley ML, Urban R, Aziz N, Chen H, Etemad-Moghadam B, Yin P;
XX
DR WPI; 2002-195801/25.
XX
PT Novel nucleic acid encoding fusion protein comprising alpha-melanocyte
PT stimulating hormone concatamer or its analog, for treating inflammatory
PT or autoimmune disorders.
XX
PS Disclosure; Page 26; 89pp; English.
XX
CC The present invention relates to a nucleic acid comprising a sequence
CC encoding a fusion polypeptide having an alpha-melanocyte stimulating
CC hormone (MSH) concatamer. The sequences are useful for treating an
CC individual suffering from, or at risk of, a disorder of the immune system
CC e.g. inflammatory disorder or autoimmune disorder, including rheumatoid
CC arthritis, asthma, sepsis, cirrhosis, dermatitis, psoriasis, contact
CC hypersensitivity, inflammatory bowel disease, autoimmune encephalitis,
CC multiple sclerosis, diabetes, lupus, uveitis and coeliac disease. The
CC present sequence is a peptide described in the exemplification of the
CC invention
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPMFLLSRNTGEVRT 15
||||||||||||||||
Db 1 TPMFLLSRNTGEVRT 15

RESULT 7
ABU96578
ID ABU96578 standard; peptide; 15 AA.
XX
AC ABU96578;
XX
DT 12-AUG-2003 (first entry)
XX
DE MHC class II associated desmoglein 3 peptide 206-220.
XX
KW Microparticle; microsphere; polynucleotide delivery; phagocytic cell;
KW tumour; viral infection; bacterial infection; fungal infection;
KW protozoan infection; gene therapy; major histocompatability complex;
KW MHC class II.
XX
OS Unidentified.
XX
PN US2002182258-A1.
XX
PD 05-DEC-2002.
XX
PF 18-JUL-2001; 2001US-00909460.
XX
PR 22-JAN-1997; 97US-0035983P.
PR 06-JAN-1998; 98US-00003253.
PR 22-JAN-1998; 98WO-US001499.
PR 11-MAR-1999; 99US-00266463.
PR 27-MAY-1999; 99US-00321346.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Lunsford LB, Putnam D, Hedley ML;
XX
DR WPI; 2003-438782/41.
XX
PT Microparticles, useful as vehicles for delivery of polynucleotides to
PT phagocytic cells, comprises polymeric matrix, lipid, and nucleic acid
PT molecule.
XX
PS Disclosure; Page 3; 37pp; English.
XX
CC The invention relates to a microparticle (microsphere) less than 20
CC microns in diameter that comprises: (1) a polymeric matrix; (2) a lipid;
CC and (3) a nucleic acid molecule. The microparticle is not encapsulated in
CC a liposome and the microparticle does not comprise a cell. The
CC microparticles are used as vehicles for the delivery of polynucleotides
CC into phagocytic cells. The microparticles can be used to express antigens
CC to treat tumour cells or viral, bacterial, fungal or protozoan
CC infections. The microparticles can be made without adversely affecting
CC nucleic acid integrity. The present sequence represents the amino acid
CC sequence of a major histocompatability complex, MHC, class II associated
CC peptide
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPMFLLSRNTGEVRT 15
||||||||||||||||
Db 1 TPMFLLSRNTGEVRT 15

RESULT 8
ADQ14319
ID ADQ14319 standard; peptide; 15 AA.
XX
AC ADQ14319;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human desmoglein 3 (DSG3) antigenic determinant (residues 206-220).
XX
KW Notch signalling; modulator; antigen; antigenic determinant;
KW immunomodulator; immune disorder; immune response; immune tolerance;
KW peripheral T-cell activation; regulatory T-cell; T reg; tumour; cancer;
KW autoimmune disorder; allergy; transplant rejection; immunosuppressive;
KW cytostatic; antiallergic; vaccine; human; autoantigen; desmoglein 3;
KW DSG3; pemphigus; MHC class II peptide; major histocompatibility complex;
KW antigenic determinant; epitope.
XX
OS Homo sapiens.
XX
PN WO2004060262-A2.
XX
PD 22-JUL-2004.
XX
PF 07-JAN-2004; 2004WO-GB000046.
XX
PR 07-JAN-2003; 2003GB-00000234.
PR 23-JAN-2003; 2003GB-00001510.
PR 23-JAN-2003; 2003GB-00001512.
PR 23-JAN-2003; 2003GB-00001513.
PR 23-JAN-2003; 2003GB-00001515.
PR 23-JAN-2003; 2003GB-00001518.
PR 23-JAN-2003; 2003GB-00001519.
PR 23-JAN-2003; 2003GB-00001521.
PR 23-JAN-2003; 2003GB-00001522.
PR 23-JAN-2003; 2003GB-00001524.

PR 23-JAN-2003; 2003GB-00001526.
PR 23-JAN-2003; 2003GB-00001527.
PR 23-JAN-2003; 2003GB-00001529.
PR 22-MAR-2003; 2003GB-00006621.
PR 04-APR-2003; 2003WO-GB001525.
PR 24-MAY-2003; 2003GB-00012062.
PR 01-AUG-2003; 2003WO-GB003285.
PR 03-OCT-2003; 2003GB-00023130.
XX
PA (LORA-) LORANTIS LTD.
XX
PI Bodmer MW, Briend ECP, Champion BR, Lennard AC, Mckenzie GJ;
PI Tugal T, Ward GA, Young LL;
XX
DR WPI; 2004-534298/51.
XX
PT New product for modulating the immune system, comprises a pharmaceutical
PT support matrix bearing modulators of Notch signaling, and an antigen or
PT antigenic determinant, or a polynucleotide coding for the antigen or
PT determinant.
XX
PS Disclosure; Page 126; 294pp; English.
XX
CC The invention relates to a product comprising (1) a pharmaceutical
CC support matrix for in vivo administration bearing modulators of Notch
CC signalling (especially a Delta or Serrate/Jagged protein or fragment or
CC homologue thereof); and (2) an antigen or antigenic determinant, or
CC polynucleotide encoding an antigen or antigenic determinant. The product
CC acts as a combined preparation for modulation of the immune system or for
CC modulation of an immune response to the antigen or antigenic determinant.
CC The invention also relates to a pharmaceutical composition comprising the
CC product; a particle with a maximum linear dimension of less than 500
CC (preferably 30-70) nm having several bound modulators of Notch signalling
CC ; a method of modulating Notch signalling; methods of treating an immune
CC disorder, for reducing an immune response, for promoting immune tolerance
CC in a mammal; and a method for increasing an immune response to a tumour
CC or pathogen antigen or its antigenic determinant in a mammal. The
CC composition and methods are useful for modulating peripheral T-cell
CC activation, for generating regulatory T-cells (T regs), for reducing an
CC immune response to an antigen or antigenic determinant, for promoting
CC immune tolerance to an antigen or antigenic determinant, or for treating
CC tumours, autoimmune disease, allergies or transplant rejection. The
CC present sequence represents an antigenic determinant derived from a human
CC autoantigen which may be used in a product of the invention.
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPMFLLSRNTGEVRT 15
|||||||||||||||
Db 1 TPMFLLSRNTGEVRT 15

RESULT 9
ADR41701
ID ADR41701 standard; peptide; 15 AA.
XX
AC ADR41701;
XX
DT 21-OCT-2004 (first entry)
XX
DE Desmoglein 3 (DSG3) (aa 206-220), class II MHC associated autoantigen.
XX
KW DSG3; Desmoglein 3; autoantigen; notch signalling pathway;
KW autoimmune disorder; bystander effect; suppression; DSL domain;
KW EGF domain; Goodpasture's disease; Wegener's granulomatosis; anaemia;
KW thrombocytopenia; gastritis; hepatitis; vasculitis; scleroderma;
KW myositis; arthritis; Systemic Lupus Erythematosus; SLE;
KW Sjogren's syndrome; hepatic fibrosis; liver cirrhosis; thyroiditis;
KW dermatitis; placental dysfunction; eclampsia;
KW inflammatory related gynaecological disease; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; Huntington disease;
KW encephalitis; psychiatric disorder; Down's syndrome; stroke; exogenous;
KW bystander antigen; multiple sclerosis; delta serrate lag;
KW inflammatory bowel disease; notch receptor.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO2004064863-A1.
XX
PD 05-AUG-2004.
XX
PF 23-JAN-2004; 2004WO-GB000263.
XX
PR 23-JAN-2003; 2003GB-00001510.
PR 23-JAN-2003; 2003GB-00001512.
PR 23-JAN-2003; 2003GB-00001513.
PR 23-JAN-2003; 2003GB-00001515.
PR 23-JAN-2003; 2003GB-00001518.
PR 23-JAN-2003; 2003GB-00001519.
PR 23-JAN-2003; 2003GB-00001521.
PR 23-JAN-2003; 2003GB-00001522.
PR 23-JAN-2003; 2003GB-00001524.
PR 23-JAN-2003; 2003GB-00001526.
PR 23-JAN-2003; 2003GB-00001527.
PR 23-JAN-2003; 2003GB-00001529.
PR 04-APR-2003; 2003WO-GB001525.
PR 24-MAY-2003; 2003GB-00012062.
PR 01-AUG-2003; 2003WO-GB003285.
PR 03-OCT-2003; 2003GB-00023130.
PR 07-JAN-2004; 2004WO-GB000046.
XX
PA (LORA-) LORANTIS LTD.
XX
PI Champion BR, Ragno S, Young LL;
XX
DR WPI; 2004-562091/54.
XX
PT New product having a modulator of the Notch signaling pathway, useful for
PT modulating an immune response in autoimmune disorders, such as anemia,
PT gastritis, hepatitis, scleroderma and myositis.
XX
PS Disclosure; Page 66; 244pp; English.
XX
CC The invention relates to the modulation of immune function through a
CC notch signalling pathway for the prevention of autoimmune diseases. It
CC has been found that the notch signalling pathway provides a bystander
CC effect or bystander suppression effect, which can be used in a wide
CC variety of ways to suppress unwanted immune responses in immune diseases
CC and disorders. Autoimmune diseases are characterised by immune responses
CC that are directed against self antigens. T lymphocytes are activated upon
CC recognition of a self antigen and/or a foreign antigen as a complex with
CC self major histocompatibility complex (MHC) gene products on the surface
CC of antigen presenting cells (APC). The invention provides the method of
CC modulating of an immune response, modulator information and a
CC pharmaceutical kit for suppression of an immune response. The modulator
CC of the notch signalling pathway is an agent which activates the notch
CC receptor or a polynucleotide which codes for such an agent. It comprises
CC a protein or polypeptide comprising a notch ligand DSL (delta serrate
CC lag) domain, notch ligand EGF domain, optionally all or part of a notch
CC ligand N terminal domain, and optionally one or more heterologous amino
CC acid or a polynucleotide sequences. The modulator can be a fusion protein
CC comprising a segment of a notch ligand extracellular domain and an
CC immunoglobulin Fc segment. The disorders include Goodpasture's disease,
CC Wegener's granulomatosis, autoimmune anaemia, thrombocytopenia,
CC gastritis, autoimmune hepatitis, inflammatory bowel disease, autoimmune
CC vasculitis, scleroderma, myositis, autoimmune arthritis, Systemic Lupus
CC Erythematosus (SLE) or Sjogren's syndrome, hepatic fibrosis, liver
CC cirrhosis, thyroiditis, dermatitis, placental dysfunction, eclampsia,
CC inflammatory related gynaecological diseases, neurodegenerative disorders
CC (such as Alzheimer's disease, Parkinson's disease, Huntington disease)
CC encephalitis, psychiatric disorders, Down's syndrome, stroke, multiple

```
CC   sclerosis, etc. The invention discloses a method for generating immune
CC   suppression at a disease locus by administering an exogenous antigen. It
CC   also provides the use of modulator or activator of notch signalling in
CC   simultaneous, separate or sequential combination with a bystander antigen
CC   or antigenic determinant for reducing an immune response to a target
CC   antigen. The presented protein sequence is the desmoglein 3 (DSG3) (aa
CC   206-220), class II MHC associated autoantigen.
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 76;  DB 8;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPMFLLSRNTGEVRT 15
              |||||||||||||||
Db          1 TPMFLLSRNTGEVRT 15


RESULT 10
ADS14311
ID   ADS14311 standard; peptide; 15 AA.
XX
AC   ADS14311;
XX
DT   16-DEC-2004  (first entry)
XX
DE   Desmoglein 3 antigenic peptide residues 206-220.
XX
KW   Cytostatic; Immunosuppressive; Antidiabetic; Neuroprotective;
KW   Antiarthritic; Antirheumatic; Antiallergic; Vaccine; Notch signaling;
KW   Notch; Notch ligand; Delta protein; Serrate protein; Jagged protein;
KW   multiple sclerosis; rheumatoid arthritis; diabetes; allergy;
KW   immune disorder; autoimmune disease; graft rejection; cancer;
KW   organ transplant; desmoglein 3.
XX
OS   Unidentified.
XX
PN   WO2004083372-A2.
XX
PD   30-SEP-2004.
XX
PF   22-MAR-2004; 2004WO-GB001229.
XX
PR   21-MAR-2003; 2003GB-00006582.
PR   21-MAR-2003; 2003GB-00006583.
PR   22-MAR-2003; 2003GB-00006621.
PR   22-MAR-2003; 2003GB-00006622.
PR   22-MAR-2003; 2003GB-00006624.
PR   22-MAR-2003; 2003GB-00006626.
PR   22-MAR-2003; 2003GB-00006640.
PR   22-MAR-2003; 2003GB-00006644.
PR   22-MAR-2003; 2003GB-00006650.
PR   22-MAR-2003; 2003GB-00006651.
PR   22-MAR-2003; 2003GB-00006654.
XX
PA   (LORA-) LORANTIS LTD.
XX
PI   Champion BR,  Ragno S;
XX
DR   WPI; 2004-709927/69.
XX
PT   Particle capable of being inserted into or taken up by cell useful for
PT   modulating immune response to antigen in subject, comprises
PT   polynucleotide coding for modulator of Notch signaling, and
PT   polynucleotide coding for antigen.
XX
PS   Disclosure; Page 119; 278pp; English.
XX
CC   The present invention relates to a particle (I) capable of being inserted
CC   into or taken up by a cell comprising (i) a polynucleotide coding for a
CC   modulator of Notch signaling, and (ii) a polynucleotide coding for an
CC   antigen or antigenic determinant. The first polynucleotide sequence codes
CC   for a Notch ligand such as a Delta or Serrate/Jagged protein or its
CC   fragment, derivative, homologue, analogue or allelic variant, or for a
CC   protein which comprises a Notch ligand DSL domain and at least one Notch
CC   ligand EGF-like domain and optionally a membrane binding or transmembrane
CC   domain. The first and second sequences are operably linked to one or more
CC   promoters or enhancers or polyadenylation sequences. The antigen or
CC   antigenic determinant is an allergen, autoantigen, Major
CC   Histocompatibility complex (MHC) (transplant) antigen, pathogen antigen,
CC   tumour antigen or their antigenic determinant. (I) is useful for
CC   modulating an immune response to an antigen in a subject. Pharmaceutical
CC   compositions comprising (I) are useful for treating or preventing
CC   conditions mediated by T cells, such as multiple sclerosis, rheumatoid
CC   arthritis, diabetes, allergy, for treating immune disorders such as
CC   autoimmune diseases of graft rejection such as allograft rejection,
CC   treating cancer and organ transplants. The present sequence is a
CC   desmoglein 3 antigenic peptide (a class II MHC-associated autoantigen
CC   peptide), which can be used as an antigen to prepare the particle of the
CC   invention.
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 76;  DB 8;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TPMFLLSRNTGEVRT 15
              |||||||||||||||
Db          1 TPMFLLSRNTGEVRT 15


RESULT 11
AAY63823
ID   AAY63823 standard; peptide; 11 AA.
XX
AC   AAY63823;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Desmoglein cell adhesion recognition cyclic peptide SEQ ID NO:3275.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
```

```
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 90; Page 209; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             56.6%;  Score 43;  DB 3;  Length 11;
  Best Local Similarity   70.0%;  Pred. No. 0.88;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MFLLSRNTGE 12
              ||:::|||||
Db          2 MFIINRNTGE 11


RESULT 12
ABB45875
ID   ABB45875 standard; peptide; 11 AA.
XX
AC   ABB45875;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Desmoglein CAR sequence cyclic peptide SEQ ID NO 619.
XX
KW   Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;
KW   cytostatic; antiapoptotic; wound healing; reduce scar tissue; skin graft;
KW   organ implant; autoimmune blistering disorder; cancer; apoptosis; cyclic.
XX
OS   Synthetic.
XX
PN   WO200172956-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-IB001400.
XX
PR   27-MAR-2000; 2000US-00535852.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Symonds JM,  Gour BJ;
XX
DR   WPI; 2002-025778/03.
XX
PT   Modulating agents for inhibiting or enhancing desmosomal cadherin
PT   mediated cell adhesion, useful for facilitating wound healing and/or
PT   reducing scar tissue, treating cancer and inducing apoptosis.
XX
PS   Claim 18; Page 99; 127pp; English.
XX
CC   The invention relates to modulating agents for inhibiting or enhancing
CC   desmosomal cadherin mediated cell adhesion, comprising a modulating agent
CC   comprising a desmosomal cadherin cell adhesion recognition CAR sequence
CC   (ABB45341-ABB47262), a non-peptide mimetic of a desmosomal cadherin CAR
CC   sequence, a substance such as an antibody or antigen-binding fragment
CC   that specifically binds a desmosomal cadherin CAR sequence and/or a
CC   polynucleotide encoding a polypeptide that comprises a desmosomal
CC   cadherin CAR sequence or analogue. The modulating agents have
CC   immunosuppressive, cytostatic and antiapoptotic activity and are used to
CC   facilitate wound healing and/or reduce scar tissue, for enhancing
CC   adhesion of foreign tissue implants (e.g. skin graft or organ implant),
CC   treating an autoimmune blistering disorder and to treat cancer (e.g.
CC   carcinoma, leukaemia or melanoma) and induce apoptosis
XX
SQ   Sequence 11 AA;

  Query Match             56.6%;  Score 43;  DB 5;  Length 11;
  Best Local Similarity   70.0%;  Pred. No. 0.88;
  Matches    7;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MFLLSRNTGE 12
              ||:::|||||
Db          2 MFIINRNTGE 11


RESULT 13
AAY63733
ID   AAY63733 standard; peptide; 11 AA.
XX
AC   AAY63733;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Desmoglein cell adhesion recognition cyclic peptide SEQ ID NO:3185.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
```

```
XX
PT   New cadherin modulating agents, used for modulating nonclassical cadherin
PT   -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT   arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS   Claim 90; Page 208; 252pp; English.
XX
CC   The present invention describes cadherin modulating agents (MA)
CC   comprising peptides which comprise a nonclassical cadherin cell adhesion
CC   recognition (CAR) sequence. The MAs can be used for modulating
CC   nonclassical cadherin-mediated functions. They can be used for e.g.
CC   inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC   mammal, enhancing delivery of a drug through the skin of a mammal,
CC   enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC   a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC   angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC   expressing cell, preventing or treating obesity in a mammal, stimulating
CC   blood vessel regression in a mammal, enhancing drug delivery to the
CC   central nervous system, treating a demyelinating neurological disease,
CC   increasing vasopermeability in a mammal, enhancing adhesion of
CC   nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC   a mammal, or preventing pregnancy in a mammal. They can also be used for
CC   e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC   or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC   mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC   -related macular degeneration, multiple sclerosis and diabetes. The
CC   products can also be used for detection and diagnosis and in bioreactors.
CC   AAY60592 to AAY64572 represent specifically claimed peptides, and
CC   AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC   the exemplification of the present invention
XX
SQ   Sequence 11 AA;

  Query Match             52.6%;  Score 40;  DB 3;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 3.1;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MFLLSRNTGE 12
              ||:::||||:
Db          2 MFIINRNTGD 11


RESULT 14
ABB45785
ID   ABB45785 standard; peptide; 11 AA.
XX
AC   ABB45785;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Desmoglein CAR sequence cyclic peptide SEQ ID NO 529.
XX
KW   Desmosomal cadherin; cell adhesion; CAR sequence; immunosuppressive;
KW   cytostatic; antiapoptotic; wound healing; reduce scar tissue; skin graft;
KW   organ implant; autoimmune blistering disorder; cancer; apoptosis; cyclic.
XX
OS   Synthetic.
XX
PN   WO200172956-A2.
XX
PD   04-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-IB001400.
XX
PR   27-MAR-2000; 2000US-00535852.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Symonds JM,  Gour BJ;
XX
DR   WPI; 2002-025778/03.
XX
PT   Modulating agents for inhibiting or enhancing desmosomal cadherin
PT   mediated cell adhesion, useful for facilitating wound healing and/or
PT   reducing scar tissue, treating cancer and inducing apoptosis.
XX
PS   Claim 18; Page 98; 127pp; English.
XX
CC   The invention relates to modulating agents for inhibiting or enhancing
CC   desmosomal cadherin mediated cell adhesion, comprising a modulating agent
CC   comprising a desmosomal cadherin cell adhesion recognition CAR sequence
CC   (ABB45341-ABB47262), a non-peptide mimetic of a desmosomal cadherin CAR
CC   sequence, a substance such as an antibody or antigen-binding fragment
CC   that specifically binds a desmosomal cadherin CAR sequence and/or a
CC   polynucleotide encoding a polypeptide that comprises a desmosomal
CC   cadherin CAR sequence or analogue. The modulating agents have
CC   immunosuppressive, cytostatic and antiapoptotic activity and are used to
CC   facilitate wound healing and/or reduce scar tissue, for enhancing
CC   adhesion of foreign tissue implants (e.g. skin graft or organ implant),
CC   treating an autoimmune blistering disorder and to treat cancer (e.g.
CC   carcinoma, leukaemia or melanoma) and induce apoptosis
XX
SQ   Sequence 11 AA;

  Query Match             52.6%;  Score 40;  DB 5;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 3.1;
  Matches    6;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 MFLLSRNTGE 12
              ||:::||||:
Db          2 MFIINRNTGD 11


RESULT 15
AAY63778
ID   AAY63778 standard; peptide; 11 AA.
XX
AC   AAY63778;
XX
DT   02-MAR-2000  (first entry)
XX
DE   Desmoglein cell adhesion recognition cyclic peptide SEQ ID NO:3230.
XX
KW   Modulation; nonclassical cadherin mediated cell adhesion; CAR;
KW   inhibition; cadherin extracellular domain; cell adhesion recognition;
KW   OB-cadherin; cadherin-5; cadherin-6; cadherin-7; cadherin-8; cadherin-12;
KW   cadherin-14; cadherin-15; T-cadherin; PB-cadherin;
KW   cadherin related neuronal receptor; LI-cadherin; protocadherin;
KW   desmoglein; desmocollin; calcium binding; cancer; tumour; obesity;
KW   rheumatoid arthritis; multiple sclerosis; diabetes; metastasis;
KW   neurological disease; cyclic.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1. .11
FT                   /note= "the terminal residues are condensed with each
FT                   other to form a cyclic peptide"
XX
PN   WO9957149-A2.
XX
PD   11-NOV-1999.
XX
PF   05-MAY-1999;   99WO-CA000363.
XX
PR   05-MAY-1998;   98US-00073040.
PR   06-NOV-1998;   98US-00187859.
PR   20-JAN-1999;   99US-00234395.
PR   08-MAR-1999;   99US-00264516.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuk OW,  Gour BJ,  Byers S;
XX
DR   WPI; 2000-038791/03.
```

XX
PT New cadherin modulating agents, used for modulating nonclassical cadherin
PT -mediated functions for treating e.g. cancers, obesity, rheumatoid
PT arthritis, multiple sclerosis, diabetes or a neurological disease.
XX
PS Claim 90; Page 209; 252pp; English.
XX
CC The present invention describes cadherin modulating agents (MA)
CC comprising peptides which comprise a nonclassical cadherin cell adhesion
CC recognition (CAR) sequence. The MAs can be used for modulating
CC nonclassical cadherin-mediated functions. They can be used for e.g.
CC inhibiting adhesion of nonclassical-cadherin expressing cells in a
CC mammal, enhancing delivery of a drug through the skin of a mammal,
CC enhancing delivery of a drug to a tumour in a mammal, treating cancer in
CC a mammal, inhibiting metastasis of a cancer in a mammal, inhibiting
CC angiogenesis in a mammal, inducing apoptosis in a nonclassical cadherin-
CC expressing cell, preventing or treating obesity in a mammal, stimulating
CC blood vessel regression in a mammal, enhancing drug delivery to the
CC central nervous system, treating a demyelinating neurological disease,
CC increasing vasopermeability in a mammal, enhancing adhesion of
CC nonclassical cadherin-expressing cells, inhibiting synaptic stability in
CC a mammal, or preventing pregnancy in a mammal. They can also be used for
CC e.g. enhancing or directing neurite outgrowth, facilitating wound healing
CC or reducing scar tissue, or enhancing adhesion of foreign tissue in a
CC mammal. They can also be used for treating e.g. psoriasis, arthritis, age
CC -related macular degeneration, multiple sclerosis and diabetes. The
CC products can also be used for detection and diagnosis and in bioreactors.
CC AAY60592 to AAY64572 represent specifically claimed peptides, and
CC AAY64573 to AAY64643 and AAZ33183 to AAZ33186 represent sequences used in
CC the exemplification of the present invention
XX
SQ Sequence 11 AA;

Query Match 51.3%; Score 39; DB 3; Length 11;
Best Local Similarity 60.0%; Pred. No. 4.6;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 MFLLSRNTGE 12
              ||:::||||:
Db          2 MFIINRNTGK 11
```

Search completed: February 22, 2005, 09:24:36
Job time : 66.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-5
Perfect score: 88
Sequence: 1 CECNIKVKDVNDNFP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	31.8	15	2	Q9S0V1	Q9s0v1 nitrogen fi
2	26	29.5	9	2	Q48686	Q48686 lactococcus
3	25	28.4	12	2	Q9ICE5	Q9ice5 human immun
4	25	28.4	15	2	Q7S362	Q7s362 neurospora
5	24	27.3	8	1	PLP_BRANA	P81707 brassica na
6	24	27.3	11	1	CSI5_BACSU	P81095 bacillus su
7	24	27.3	13	2	Q9KHJ4	Q9khj4 caulobacter
8	24	27.3	14	2	Q661E1	Q661e1 borrelia ga
9	24	27.3	15	1	VSP3_AGKHP	P80899 agkistrodon
10	24	27.3	15	2	Q9PRW2	Q9prw2 crotalus at
11	24	27.3	15	2	Q9PRW3	Q9prw3 crotalus at
12	23	26.1	14	2	Q7X9S0	Q7x9s0 malus domes
13	23	26.1	15	2	Q7R923	Q7r923 plasmodium
14	23	26.1	15	2	Q9TRP2	Q9trp2 sus scrofa
15	23	26.1	15	2	Q9QV01	Q9qv01 mus sp. . 1
16	22	25.0	8	2	Q9N6M5	Q9n6m5 toxoplasma
17	22	25.0	11	2	Q48933	Q48933 mycobacteri
18	22	25.0	11	2	Q79C20	Q79c20 mycobacteri
19	22	25.0	11	2	Q79C22	Q79c22 mycobacteri
20	22	25.0	11	2	Q707P1	Q707p1 rattus norv
21	22	25.0	13	2	Q7M0J8	Q7m0j8 planobispor
22	22	25.0	14	1	ANS4_ANISI	P83885 anisakis si
23	22	25.0	14	2	Q6JVP2	Q6jvp2 otopappus e
24	22	25.0	14	2	Q6JVP4	Q6jvp4 jefea pring
25	22	25.0	14	2	Q6JVQ6	Q6jvq6 angelphytum
26	22	25.0	15	2	Q9BXX4	Q9bxx4 homo sapien
27	22	25.0	15	2	O19468	O19468 mus musculu
28	22	25.0	15	2	Q6LCH6	Q6lch6 mus musculu
29	21	23.9	10	2	Q9UMK9	Q9umk9 homo sapien
30	21	23.9	11	2	Q9UBM2	Q9ubm2 homo sapien
31	21	23.9	11	2	Q7M372	Q7m372 bos taurus
32	21	23.9	11	2	Q7ZDA8	Q7zda8 human immun
33	21	23.9	14	2	Q6JVP5	Q6jvp5 elaphandra
34	21	23.9	14	2	Q6JVP7	Q6jvp7 dimerostemm
35	21	23.9	14	2	Q6JVQ1	Q6jvq1 blainvillea
36	21	23.9	14	2	Q6JVQ2	Q6jvq2 baltimora r
37	21	23.9	14	2	Q6JVQ4	Q6jvq4 angelphytum
38	21	23.9	15	2	Q6JVP9	Q6jvp9 dimerostemm
39	21	23.9	15	2	Q08936	Q08936 nicotiana t
40	21	23.9	15	2	Q7M279	Q7m279 oryza sativ
41	21	23.9	15	2	Q6LCW3	Q6lcw3 mus musculu
42	20.5	23.3	15	2	O07478	O07478 streptococc
43	20	22.7	8	2	Q6UA69	Q6ua69 carassius c
44	20	22.7	9	2	Q6UVK2	Q6uvk2 malus domes
45	20	22.7	10	2	Q7RH60	Q7rh60 plasmodium

ALIGNMENTS

RESULT 1
Q9S0V1
ID Q9S0V1 PRELIMINARY; PRT; 15 AA.
AC Q9S0V1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Dinitrogenase beta subunit (Fragment).
GN Name=anfK;
OS nitrogen fixing bacterium ANFK33.
OC Bacteria.
OX NCBI_TaxID=96027;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20011232; PubMed=10543806;
RA Noda S., Ohkuma M., Usami R., Horikoshi K., Kudo T.;
RT "Culture-independent characterization of gene resposible for nitrogen
RT fixation in the symbiotic microbial community in the gut of the
RT termite Neotermes koshunensis.";
RL Appl. Environ. Microbiol. 65:4935-4942(1999).
DR EMBL; AB027751; BAA86283.1; -.
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1719 MW; 4D64BBE1338D8A9A CRC64;

Query Match 31.8%; Score 28; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          3 CNIKVKD 9
              | :|:||
Db          3 CELKLKD 9
```

RESULT 2
Q48686
ID Q48686 PRELIMINARY; PRT; 9 AA.
AC Q48686;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Streptococcus cremoris promoter 23 DNA. (Fragment).
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.
OX NCBI_TaxID=1359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88105390; PubMed=2447829;
RA van der Vossen J.M., van der Lelie D., Venema G.;
RT "Isolation and characterization of Streptococcus cremoris Wg2-specific
RT promoters.";
RL Appl. Environ. Microbiol. 53:2452-2457(1987).
DR EMBL; M24763; AAA74720.1; -.
FT NON_TER 9 9
SQ SEQUENCE 9 AA; 1080 MW; 5AF3A44AA4469443 CRC64;

```
  Query Match             29.5%;  Score 26;  DB 2;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 VKDVNDN 13
              :|::|||
Db          1 MKNMNDN 7


RESULT 3
Q9ICE5
ID   Q9ICE5      PRELIMINARY;      PRT;    12 AA.
AC   Q9ICE5;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Envelope glycoprotein.
GN   Name=env;
OS   Human immunodeficiency virus 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX   NCBI_TaxID=11676;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Choroid plexus;
RA   Gartner S., Liu Y., Tang X.P., McArthur J.C., Scott J.;
RT   "Analysis of human immunodeficiency virus type 1 gp160 sequences from
RT   a patient with HIV dementia: evidence for monocyte trafficking into
RT   brain.";
RL   J. Neurovirol. 0:0-0(2000).
DR   EMBL; AF217155; AAF75497.1; -.
DR   EMBL; AF217153; AAF75495.1; -.
DR   EMBL; AF217154; AAF75496.1; -.
DR   GO; GO:0019031; C:viral envelope; IEA.
KW   Envelope protein.
SQ   SEQUENCE   12 AA;  1636 MW;  7ED6A2917A24005B CRC64;

  Query Match             28.4%;  Score 25;  DB 2;  Length 12;
  Best Local Similarity   30.0%;  Pred. No. 3.9e+03;
  Matches    3;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IKVKDVNDNF 14
              ::||::  |:
Db          1 MRVKEIRKNY 10


RESULT 4
Q7S362
ID   Q7S362      PRELIMINARY;      PRT;    15 AA.
AC   Q7S362;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Predicted protein.
GN   Name=NCU08588.1;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OR74A;
RA   Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA   Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA   Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA   Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA   Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA   Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA   Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA   Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA   Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA   Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA   DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA   Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA   Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA   Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT   "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL   Nature 0:0-0(2003).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABX01000413; EAA29870.1; -.
SQ   SEQUENCE   15 AA;  1753 MW;  44BAA4C25D52DA9A CRC64;

  Query Match             28.4%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 4.8e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 VKDVNDN 13
              :||:| |
Db          7 IKDINLN 13


RESULT 5
PLP_BRANA
ID   PLP_BRANA      STANDARD;      PRT;     8 AA.
AC   P81707;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Plastidial lipid-associated protein (Fragment).
OS   Brassica napus (Rape).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids II; Brassicales; Brassicaceae; Brassica.
OX   NCBI_TaxID=3708;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=cv. Topaz; TISSUE=Tapetum;
RX   MEDLINE=99349136; PubMed=10420651;
RA   Hernandez-Pinzon I., Ross J.H.E., Barnes K.A., Damant A.P.,
RA   Murphy D.J.;
RT   "Composition and role of tapetal lipid bodies in the biogenesis of the
RT   pollen coat of Brassica napus.";
RL   Planta 208:588-598(1999).
CC   -!- FUNCTION: May play a structural role in the elaioplast, a tapetum-
CC       specific plastidial lipid organelle.
CC   -!- TISSUE SPECIFICITY: Tapetum of anthers.
KW   Direct protein sequencing.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  989 MW;  9D7B1AA452CAA042 CRC64;

  Query Match             27.3%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   62.5%;  Pred. No. 1.6e+06;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 VKDVNDNF 14
              | |||| :
Db          1 VIDVNDEW 8


RESULT 6
CSI5_BACSU
ID   CSI5_BACSU      STANDARD;      PRT;     11 AA.
AC   P81095;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Cold shock protein CSI5 (11 kDa cold shock protein) (Fragment).
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=168 / JH642;
RA   Graumann P.L., Schmid R., Marahiel M.A.;
```

```
RL   Submitted (OCT-1997) to Swiss-Prot.
RN   [2]
RP   CHARACTERIZATION.
RC   STRAIN=168 / JH642;
RX   MEDLINE=96345629; PubMed=8755892;
RA   Graumann P., Schroeder K., Schmid R., Marahiel M.A.;
RT   "Cold shock stress-induced proteins in Bacillus subtilis.";
RL   J. Bacteriol. 178:4611-4619(1996).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- INDUCTION: In response to low temperature.
CC   -!- CAUTION: Could not be found in the genome of B.subtilis 168.
KW   Direct protein sequencing.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1360 MW;  15F6ECEE6322C330 CRC64;

  Query Match             27.3%;  Score 24;  DB 1;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 5.1e+03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 NIKVK 8
              |||||
Db          3 NIKVK 7


RESULT 7
Q9KHJ4
ID   Q9KHJ4      PRELIMINARY;      PRT;    13 AA.
AC   Q9KHJ4;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Proteolysis tag (Fragment).
OS   Caulobacter crescentus.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC   Caulobacteraceae; Caulobacter.
OX   NCBI_TaxID=155892;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20345063; PubMed=10884408; DOI=10.1073/pnas.97.14.7778;
RA   Keiler K.C., Shapiro L., Williams K.P.;
RT   "tmRNAs that encode proteolysis-inducing tags are found in all known
RT   bacterial genomes: A two-piece tmRNA functions in Caulobacter.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:7778-7783(2000).
DR   EMBL; AF255738; AAF87998.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1368 MW;  CE5F60C57FCE1B1D CRC64;

  Query Match             27.3%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 6e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 NDNF 14
              ||||
Db          2 NDNF 5


RESULT 8
Q661E1
ID   Q661E1      PRELIMINARY;      PRT;    14 AA.
AC   Q661E1;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=BG0491;
OS   Borrelia garinii PBi.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=290434;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PBi;
RA   Gloeckner G., Lehmann R., Romualdi A., Pradella S.,
RA   Schulte-Spechtel U., Wilske B., Suehnel J., Platzer M.;
RT   "Comparative analysis of the Borrelia garinii genome.";
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; CP000013; AAU07330.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   14 AA;  1712 MW;  DD3944F0FB199202 CRC64;

  Query Match             27.3%;  Score 24;  DB 2;  Length 14;
  Best Local Similarity   57.1%;  Pred. No. 6.5e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 KDVNDNF 14
              ||| | :
Db          5 KDVEDKY 11


RESULT 9
VSP3_AGKHP
ID   VSP3_AGKHP     STANDARD;      PRT;    15 AA.
AC   P80899;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Ancrod (EC 3.4.21.74) (Venombin A) (Protein C activator) (ACC-C)
DE   (Fragment).
OS   Agkistrodon halys pallas (Chinese water mocassin) (Gloydius halys
OS   pallas).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Gloydius.
OX   NCBI_TaxID=8714;
RN   [1]
RP   SEQUENCE, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC   TISSUE=Venom;
RA   Hong S.;
RL   Submitted (FEB-1997) to Swiss-Prot.
CC   -!- FUNCTION: Thrombin-like snake venom serine protease. Activates
CC       protein C (By similarity).
CC   -!- CATALYTIC ACTIVITY: Selective cleavage of Arg-|-Xaa bond in
CC       fibrinogen, to form fibrin, and release fibrinopeptide A. The
CC       specificity of further degradation of fibrinogen varies with
CC       species origin of the enzyme.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Snake venom
CC       subfamily.
DR   InterPro; IPR001254; Peptidase_S1.
DR   PROSITE; PS50240; TRYPSIN_DOM; PARTIAL.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1642 MW;  03EFE10227CD8CDA CRC64;

  Query Match             27.3%;  Score 24;  DB 1;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ECNI 5
              ||||
Db          6 ECNI 9


RESULT 10
Q9PRW2
ID   Q9PRW2      PRELIMINARY;      PRT;    15 AA.
AC   Q9PRW2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Alpha-FIBRINOGENASE isoform A3 (Fragment).
OS   Crotalus atrox (Western diamondback rattlesnake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
```

```
OC   Viperidae; Crotalinae; Crotalus.
OX   NCBI_TaxID=8730;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94296418; PubMed=8024586;
RA   Hung C.C., Chiou S.H.;
RT   "Isolation of multiple isoforms of alpha-fibrinogenase from the
RT   Western diamondback rattlesnake, Crotalus atrox: N-terminal sequence
RT   homology with ancrod, an antithrombotic agent from Malayan viper.";
RL   Biochem. Biophys. Res. Commun. 201:1414-1423(1994).
SQ   SEQUENCE   15 AA;  1656 MW;  03EFE10227D52FDA CRC64;

  Query Match             27.3%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ECNI 5
              ||||
Db          6 ECNI 9


RESULT 11
Q9PRW3
ID   Q9PRW3      PRELIMINARY;      PRT;    15 AA.
AC   Q9PRW3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Alpha-FIBRINOGENASE isoform A2 (Fragment).
OS   Crotalus atrox (Western diamondback rattlesnake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Crotalus.
OX   NCBI_TaxID=8730;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94296418; PubMed=8024586;
RA   Hung C.C., Chiou S.H.;
RT   "Isolation of multiple isoforms of alpha-fibrinogenase from the
RT   Western diamondback rattlesnake, Crotalus atrox: N-terminal sequence
RT   homology with ancrod, an antithrombotic agent from Malayan viper.";
RL   Biochem. Biophys. Res. Commun. 201:1414-1423(1994).
DR   PIR; PC2215; PC2215.
SQ   SEQUENCE   15 AA;  1640 MW;  03EFE10227CA12DA CRC64;

  Query Match             27.3%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 ECNI 5
              ||||
Db          6 ECNI 9


RESULT 12
Q7X9S0
ID   Q7X9S0      PRELIMINARY;      PRT;    14 AA.
AC   Q7X9S0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Metallothionein-like protein (Fragment).
OS   Malus domestica (Apple) (Malus sylvestris).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX   NCBI_TaxID=3750;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Fruit pulp;
RA   Pathak N., Asif M., Solomos T.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY277667; AAP34401.1; -.
DR   GO; GO:0046872; F:metal ion binding; IEA.
DR   InterPro; IPR000347; Metallthion_15p.
DR   Pfam; PF01439; Metallothio_2; 1.
DR   ProDom; PD001611; Metallthion_15p; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1503 MW;  85ED164D6EB3E9A1 CRC64;

  Query Match             26.1%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   75.0%;  Pred. No. 9.3e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 CECN 4
              | ||
Db          6 CTCN 9


RESULT 13
Q7R923
ID   Q7R923      PRELIMINARY;      PRT;    15 AA.
AC   Q7R923;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   Name=PY07046;
OS   Plasmodium yoelii yoelii.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=73239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=17XNL;
RX   PubMed=12368865; DOI=10.1038/nature01099;
RA   Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA   Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA   Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA   Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA   Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA   Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA   Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA   van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA   Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA   Carucci D.J.;
RT   "Genome sequence and comparative analysis of the model rodent malaria
RT   parasite Plasmodium yoelii yoelii.";
RL   Nature 419:512-519(2002).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABL01002504; EAA19396.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   15 AA;  1791 MW;  4A5E699572F38B6A CRC64;

  Query Match             26.1%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 1e+04;
  Matches    5;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 CNIKVKDVNDN 13
              |  :|   |||
Db          2 CAHRVIKYNDN 12


RESULT 14
Q9TRP2
ID   Q9TRP2      PRELIMINARY;      PRT;    15 AA.
AC   Q9TRP2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   LOW MR zona pellucida binding protein (Fragment).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
```

```
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92378826; PubMed=1510840;
RA   Parry R.V., Barker P.J., Jones R.;
RT   "Characterization of low Mr zona pellucida binding proteins from boar
RT   spermatozoa and seminal plasma.";
RL   Mol. Reprod. Dev. 33:108-115(1992).
DR   HSSP; P35496; 1SPP.
SQ   SEQUENCE   15 AA;  1643 MW;  7FB277857862B4BE CRC64;

  Query Match             26.1%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 1e+04;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 ECNIKVKD 9
              ||   :||
Db          8 ECGRVIKD 15


RESULT 15
Q9QV01
ID   Q9QV01      PRELIMINARY;      PRT;    15 AA.
AC   Q9QV01;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   16 kDa alpha-lactalbumin homolog (Fragment).
OS   Mus sp. .
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10095;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95365417; PubMed=7638241;
RA   Goto Y., Nagasawa H., Sasaki T., Enami J., Iguchi T.;
RT   "Biochemical changes during growth and regression of pregnancy-
RT   dependent mammary tumors of GR/A mice.";
RL   Proc. Soc. Exp. Biol. Med. 209:343-353(1995).
SQ   SEQUENCE   15 AA;  1686 MW;  D6AF74D707F8A4A6 CRC64;

  Query Match             26.1%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   30.0%;  Pred. No. 1e+04;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CECNIKVKDV 10
              |: :  :||:
Db          6 CKVSHAIKDI 15


Search completed: February 22, 2005, 09:37:53
Job time : 54.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-5
Perfect score: 88
Sequence: 1 CECNIKVKDVNDNFP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	88	100.0	15	2	AAW04845	Aaw04845 Self epit
2	45	51.1	15	6	ABR31812	Abr31812 Human can
3	45	51.1	15	6	ABR32072	Abr32072 Human can
4	45	51.1	15	6	ABR31894	Abr31894 Human can
5	45	51.1	15	6	ABR32303	Abr32303 Human can
6	45	51.1	15	6	ABR32121	Abr32121 Human can
7	45	51.1	15	6	ABR32187	Abr32187 Human can
8	44	50.0	11	2	AAW13139	Aaw13139 Human cad
9	40	45.5	15	6	ABR32064	Abr32064 Human can
10	40	45.5	15	6	ABR32177	Abr32177 Human can
11	40	45.5	15	6	ABR32247	Abr32247 Human can
12	40	45.5	15	6	ABR31714	Abr31714 Human can
13	40	45.5	15	6	ABR32176	Abr32176 Human can
14	40	45.5	15	6	ABR32292	Abr32292 Human can
15	40	45.5	15	6	ABR32186	Abr32186 Human can
16	40	45.5	15	6	ABR32178	Abr32178 Human can
17	39	44.3	10	5	AAE24278	Aae24278 Murine E-
18	38	43.2	15	6	ABR32086	Abr32086 Human can
19	38	43.2	15	6	ABR32002	Abr32002 Human can
20	38	43.2	15	6	ABR32267	Abr32267 Human can
21	38	43.2	15	6	ABR31695	Abr31695 Human can
22	38	43.2	15	6	ABR32207	Abr32207 Human can
23	38	43.2	15	6	ABR32001	Abr32001 Human can
24	37	42.0	10	6	ABR06197	Abr06197 Human can
25	36	40.9	9	6	ABR06723	Abr06723 Human can
26	36	40.9	9	6	ABR07362	Abr07362 Human can
27	36	40.9	15	4	AAG98748	Aag98748 Human cel
28	36	40.9	15	6	ABR32153	Abr32153 Human can
29	36	40.9	15	6	ABR31871	Abr31871 Human can
30	36	40.9	15	6	ABR31932	Abr31932 Human can
31	36	40.9	15	6	ABR31736	Abr31736 Human can
32	36	40.9	15	6	ABR31966	Abr31966 Human can
33	36	40.9	15	6	ABR31709	Abr31709 Human can
34	36	40.9	15	6	ABR32272	Abr32272 Human can
35	36	40.9	15	6	ABR31916	Abr31916 Human can
36	36	40.9	15	6	ABR31831	Abr31831 Human can
37	34	38.6	9	6	ABR07139	Abr07139 Human can
38	34	38.6	10	6	ABR06819	Abr06819 Human can
39	34	38.6	10	6	ABR06251	Abr06251 Human can
40	34	38.6	10	6	ABR06435	Abr06435 Human can
41	34	38.6	15	6	ABR32206	Abr32206 Human can
42	34	38.6	15	6	ABR31669	Abr31669 Human can
43	34	38.6	15	6	ABR31723	Abr31723 Human can
44	34	38.6	15	6	ABR31976	Abr31976 Human can
45	34	38.6	15	6	ABR32208	Abr32208 Human can

ALIGNMENTS

RESULT 1
AAW04845
ID AAW04845 standard; peptide; 15 AA.
XX
AC AAW04845;
XX
DT 18-FEB-1997 (first entry)
XX
DE Self epitope of desmoglein 3, implicated in autoimmune disease.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Homo sapiens.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 1; Page 40; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is derived from the human desmoglein

```
CC   3 protein (amino acids 251-265) and is implicated as a self epitope in
CC   pemphigus vulgaris. Peptides derived from the human desmoglein protein
CC   are described in AAW04841-47
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 88;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-07;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CECNIKVKDVNDNFP 15
              |||||||||||||||
Db          1 CECNIKVKDVNDNFP 15


RESULT 2
ABR31812
ID   ABR31812 standard; peptide; 15 AA.
XX
AC   ABR31812;
XX
DT   19-MAY-2003  (first entry)
XX
DE   Human cancer-related protein 109P1D4 HLA peptide #1584.
XX
KW   Human; cytostatic; vaccine; cancer; immune response; HLA;
KW   human leukocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200283921-A2.
XX
PD   24-OCT-2002.
XX
PF   10-APR-2002; 2002WO-US011654.
XX
PR   10-APR-2001; 2001US-0282739P.
PR   10-APR-2001; 2001US-0283112P.
PR   25-APR-2001; 2001US-0286630P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Challita-Eid PM,  Faris M,  Ge W,  Hubert RS;
PI   Morrison K,  Morrison RK,  Raitano AB;
XX
DR   WPI; 2003-075555/07.
XX
PT   New composition comprising a substance that modulates the structure of
PT   proteins and polynucleotides, useful for therapeutic, prognostic and
PT   diagnostic reagents for eliciting cellular or humoral immune response in
PT   cancer patients.
XX
PS   Claim 13; Page 499; 1021pp; English.
XX
CC   The present invention relates to novel human cancer-related genes and
CC   proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC   proteins are useful for eliciting a humoral or cellular immune response.
CC   The genes are useful as probes and primers for the amplification and/or
CC   detection of genes, mRNAs or their fragments, as reagents for the
CC   diagnosis and/or prognosis of cancer, as coding sequences capable of
CC   directing the expression of the protein, as tools for modulating or
CC   inhibiting the expression of genes and/or translation of transcripts, and
CC   as therapeutic agents. The proteins and peptides are useful as
CC   therapeutic, prognostic and diagnostic reagents for cancer. The present
CC   sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC   from the invention
XX
SQ   Sequence 15 AA;

  Query Match             51.1%;  Score 45;  DB 6;  Length 15;
  Best Local Similarity   81.8%;  Pred. No. 2;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IKVKDVNDNFP 15
              ||||| ||| |
Db          5 IKVKDENDNAP 15


RESULT 3
ABR32072
ID   ABR32072 standard; peptide; 15 AA.
XX
AC   ABR32072;
XX
DT   19-MAY-2003  (first entry)
XX
DE   Human cancer-related protein 109P1D4 HLA peptide #1844.
XX
KW   Human; cytostatic; vaccine; cancer; immune response; HLA;
KW   human leukocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200283921-A2.
XX
PD   24-OCT-2002.
XX
PF   10-APR-2002; 2002WO-US011654.
XX
PR   10-APR-2001; 2001US-0282739P.
PR   10-APR-2001; 2001US-0283112P.
PR   25-APR-2001; 2001US-0286630P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Challita-Eid PM,  Faris M,  Ge W,  Hubert RS;
PI   Morrison K,  Morrison RK,  Raitano AB;
XX
DR   WPI; 2003-075555/07.
XX
PT   New composition comprising a substance that modulates the structure of
PT   proteins and polynucleotides, useful for therapeutic, prognostic and
PT   diagnostic reagents for eliciting cellular or humoral immune response in
PT   cancer patients.
XX
PS   Claim 13; Page 504; 1021pp; English.
XX
CC   The present invention relates to novel human cancer-related genes and
CC   proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC   proteins are useful for eliciting a humoral or cellular immune response.
CC   The genes are useful as probes and primers for the amplification and/or
CC   detection of genes, mRNAs or their fragments, as reagents for the
CC   diagnosis and/or prognosis of cancer, as coding sequences capable of
CC   directing the expression of the protein, as tools for modulating or
CC   inhibiting the expression of genes and/or translation of transcripts, and
CC   as therapeutic agents. The proteins and peptides are useful as
CC   therapeutic, prognostic and diagnostic reagents for cancer. The present
CC   sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC   from the invention
XX
SQ   Sequence 15 AA;

  Query Match             51.1%;  Score 45;  DB 6;  Length 15;
  Best Local Similarity   81.8%;  Pred. No. 2;
  Matches    9;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IKVKDVNDNFP 15
              ||||| ||| |
Db          3 IKVKDENDNAP 13


RESULT 4
ABR31894
ID   ABR31894 standard; peptide; 15 AA.
XX
AC   ABR31894;
```

XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 109P1D4 HLA peptide #1666.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 501; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 51.1%; Score 45; DB 6; Length 15;
Best Local Similarity 81.8%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IKVKDVNDNFP 15
||||| ||| |
Db 3 IKVKDENDNAP 13

RESULT 5
ABR32303
ID ABR32303 standard; peptide; 15 AA.
XX
AC ABR32303;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 109P1D4 HLA peptide #2075.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 508; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 51.1%; Score 45; DB 6; Length 15;
Best Local Similarity 81.8%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IKVKDVNDNFP 15
||||| ||| |
Db 5 IKVKDENDNAP 15

RESULT 6
ABR32121
ID ABR32121 standard; peptide; 15 AA.
XX
AC ABR32121;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 109P1D4 HLA peptide #1893.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX

PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 505; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 51.1%; Score 45; DB 6; Length 15;
Best Local Similarity 81.8%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IKVKDVNDNFP 15
||||| ||| |
Db 5 IKVKDENDNAP 15

RESULT 7
ABR32187
ID ABR32187 standard; peptide; 15 AA.
XX
AC ABR32187;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 109P1D4 HLA peptide #1959.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 506; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 51.1%; Score 45; DB 6; Length 15;
Best Local Similarity 81.8%; Pred. No. 2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 IKVKDVNDNFP 15
||||| ||| |
Db 2 IKVKDENDNAP 12

RESULT 8
AAW13139
ID AAW13139 standard; protein; 11 AA.
XX
AC AAW13139;
XX
DT 25-MAR-2003 (revised)
DT 14-MAY-1997 (first entry)
XX
DE Human cadherin-5 antigenic epitope (residues 242-252).
XX
KW Ca2+ dependent; cell adhesion protein; cadherin; human; antibody;
KW purification; determination; epitope; tissue expression;
KW binding antagonist; calcium ion; antigen.
XX
OS Homo sapiens.
XX
PN US5597725-A.
XX
PD 28-JAN-1997.
XX
PF 26-JAN-1994; 94US-00188228.
XX
PR 17-APR-1992; 92US-00872643.
PR 19-APR-1993; 93US-00049460.
XX
PA (DOHE-) DOHENY EYE INST.
XX
PI Suzuki S;
XX
DR WPI; 1997-108328/10.
XX
PT Antibodies to cadherin proteins - useful as cadherin antagonists, etc.
XX
PS Claim 5; Col 112; 59pp; English.
XX
CC The present sequence is an antigenic epitope from human cadherin-5, which
CC is a Ca2+ dependent cell adhesion protein. Antibodies or fragments that
CC specifically bind the epitope can be used to purify the cadherin,
CC determine its tissue expression and antagonise its ligand/antiligand
CC binding activities. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 11 AA;

Query Match 50.0%; Score 44; DB 2; Length 11;

```
  Best Local Similarity   54.5%;  Pred. No. 2.1;
  Matches    6;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IKVKDVNDNFP 15
              : ::|:|||||
Db          1 VTLQDINDNFP 11


RESULT 9
ABR32064
ID   ABR32064 standard; peptide; 15 AA.
XX
AC   ABR32064;
XX
DT   19-MAY-2003  (first entry)
XX
DE   Human cancer-related protein 109P1D4 HLA peptide #1836.
XX
KW   Human; cytostatic; vaccine; cancer; immune response; HLA;
KW   human leukocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200283921-A2.
XX
PD   24-OCT-2002.
XX
PF   10-APR-2002; 2002WO-US011654.
XX
PR   10-APR-2001; 2001US-0282739P.
PR   10-APR-2001; 2001US-0283112P.
PR   25-APR-2001; 2001US-0286630P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Challita-Eid PM,  Faris M,  Ge W,  Hubert RS;
PI   Morrison K,  Morrison RK,  Raitano AB;
XX
DR   WPI; 2003-075555/07.
XX
PT   New composition comprising a substance that modulates the structure of
PT   proteins and polynucleotides, useful for therapeutic, prognostic and
PT   diagnostic reagents for eliciting cellular or humoral immune response in
PT   cancer patients.
XX
PS   Claim 13; Page 504; 1021pp; English.
XX
CC   The present invention relates to novel human cancer-related genes and
CC   proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC   proteins are useful for eliciting a humoral or cellular immune response.
CC   The genes are useful as probes and primers for the amplification and/or
CC   detection of genes, mRNAs or their fragments, as reagents for the
CC   diagnosis and/or prognosis of cancer, as coding sequences capable of
CC   directing the expression of the protein, as tools for modulating or
CC   inhibiting the expression of genes and/or translation of transcripts, and
CC   as therapeutic agents. The proteins and peptides are useful as
CC   therapeutic, prognostic and diagnostic reagents for cancer. The present
CC   sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC   from the invention
XX
SQ   Sequence 15 AA;

  Query Match             45.5%;  Score 40;  DB 6;  Length 15;
  Best Local Similarity   63.6%;  Pred. No. 13;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 IKVKDVNDNFP 15
              : | ||||| |
Db          3 VNVTDVNDNVP 13


RESULT 10
ABR32177
ID   ABR32177 standard; peptide; 15 AA.
XX
AC   ABR32177;
XX
DT   19-MAY-2003  (first entry)
XX
DE   Human cancer-related protein 109P1D4 HLA peptide #1949.
XX
KW   Human; cytostatic; vaccine; cancer; immune response; HLA;
KW   human leukocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200283921-A2.
XX
PD   24-OCT-2002.
XX
PF   10-APR-2002; 2002WO-US011654.
XX
PR   10-APR-2001; 2001US-0282739P.
PR   10-APR-2001; 2001US-0283112P.
PR   25-APR-2001; 2001US-0286630P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Challita-Eid PM,  Faris M,  Ge W,  Hubert RS;
PI   Morrison K,  Morrison RK,  Raitano AB;
XX
DR   WPI; 2003-075555/07.
XX
PT   New composition comprising a substance that modulates the structure of
PT   proteins and polynucleotides, useful for therapeutic, prognostic and
PT   diagnostic reagents for eliciting cellular or humoral immune response in
PT   cancer patients.
XX
PS   Claim 13; Page 506; 1021pp; English.
XX
CC   The present invention relates to novel human cancer-related genes and
CC   proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC   proteins are useful for eliciting a humoral or cellular immune response.
CC   The genes are useful as probes and primers for the amplification and/or
CC   detection of genes, mRNAs or their fragments, as reagents for the
CC   diagnosis and/or prognosis of cancer, as coding sequences capable of
CC   directing the expression of the protein, as tools for modulating or
CC   inhibiting the expression of genes and/or translation of transcripts, and
CC   as therapeutic agents. The proteins and peptides are useful as
CC   therapeutic, prognostic and diagnostic reagents for cancer. The present
CC   sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC   from the invention
XX
SQ   Sequence 15 AA;

  Query Match             45.5%;  Score 40;  DB 6;  Length 15;
  Best Local Similarity   63.6%;  Pred. No. 13;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 IKVKDVNDNFP 15
              : | ||||| |
Db          4 VNVTDVNDNVP 14


RESULT 11
ABR32247
ID   ABR32247 standard; peptide; 15 AA.
XX
AC   ABR32247;
XX
DT   19-MAY-2003  (first entry)
XX
DE   Human cancer-related protein 109P1D4 HLA peptide #2019.
XX
KW   Human; cytostatic; vaccine; cancer; immune response; HLA;
KW   human leukocyte antigen.
```

XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 507; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 45.5%; Score 40; DB 6; Length 15;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 IKVKDVNDNFP 15
: | ||||| |
Db 2 VNVTDVNDNVP 12

RESULT 12
ABR31714
ID ABR31714 standard; peptide; 15 AA.
XX
AC ABR31714;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 109P1D4 HLA peptide #1486.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 497; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 45.5%; Score 40; DB 6; Length 15;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 IKVKDVNDNFP 15
: | ||||| |
Db 2 VNVTDVNDNVP 12

RESULT 13
ABR32176
ID ABR32176 standard; peptide; 15 AA.
XX
AC ABR32176;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 109P1D4 HLA peptide #1948.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX

PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 506; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 45.5%; Score 40; DB 6; Length 15;
Best Local Similarity 63.6%; Pred. No. 13;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 IKVKDVNDNFP 15
: | ||||| |
Db 5 VNVTDVNDNVP 15

RESULT 14
ABR32292
ID ABR32292 standard; peptide; 15 AA.
XX
AC ABR32292;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 109P1D4 HLA peptide #2064.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 508; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 45.5%; Score 40; DB 6; Length 15;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 IKVKDVNDN 13
||||| |||
Db 6 IKVKDENDN 14

RESULT 15
ABR32186
ID ABR32186 standard; peptide; 15 AA.
XX
AC ABR32186;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 109P1D4 HLA peptide #1958.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 506; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX

SQ Sequence 15 AA;

Query Match 45.5%; Score 40; DB 6; Length 15;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          5 IKVKDVNDN 13
              ||||| |||
Db          7 IKVKDENDN 15
```

Search completed: February 22, 2005, 09:24:37
Job time : 66.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-6
Perfect score: 82
Sequence: 1 SARTLNNRYTGPYTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	32.9	9	2	S70332	endosperm protein,
2	24	29.3	15	2	A56970	GLYMA1 - soybean (
3	23	28.0	15	2	S08209	hypothetical prote
4	22	26.8	12	2	A28856	fructose-bisphosph
5	22	26.8	12	2	PH1459	T-cell receptor be
6	22	26.8	15	2	S36893	ribosomal protein
7	22	26.8	15	2	A28497	neurotensin-relate
8	21	25.6	8	2	D47393	neuropeptide calla
9	21	25.6	13	2	PH1593	Ig H chain V-D-J r
10	21	25.6	15	2	D28587	T-cell receptor be
11	21	25.6	15	2	F28587	T-cell receptor be
12	21	25.6	15	2	I53284	T-cell receptor be
13	21	25.6	15	4	I38335	hypothetical TEL/M
14	20.5	25.0	12	2	PH1462	T-cell receptor be
15	20.5	25.0	13	2	PH0799	T-cell receptor al
16	20	24.4	10	2	PH0927	T-cell receptor be
17	20	24.4	11	2	PH1375	T antigen variant
18	20	24.4	12	2	S26546	T-cell receptor be
19	20	24.4	12	2	S23168	Z protein - guinea
20	20	24.4	12	2	PH1461	T-cell receptor be
21	20	24.4	13	2	A32734	enkephalin precurs
22	20	24.4	13	2	S23372	T-cell receptor al
23	20	24.4	13	2	PH1585	Ig H chain V-D-J r
24	20	24.4	13	2	PH0783	T-cell receptor al
25	20	24.4	14	2	PH1305	Ig heavy chain DJ
26	20	24.4	15	2	PW0004	chlorophyll a/b-bi
27	20	24.4	15	2	PH1318	Ig heavy chain DJ
28	20	24.4	15	2	B45115	peptidylprolyl iso
29	19	23.2	8	2	E47393	neuropeptide calla
30	19	23.2	9	2	S39766	cardioactive pepti
31	19	23.2	9	2	S39767	cardioactive pepti
32	19	23.2	9	2	A26363	cardioactive pepti
33	19	23.2	9	2	S27233	cardioactive pepti
34	19	23.2	10	2	C61440	polygalacturonase
35	19	23.2	10	2	S07202	phyllomedulin - tw
36	19	23.2	12	2	S17869	glutathione transf
37	19	23.2	12	2	S49547	hypothetical prote
38	19	23.2	13	2	S47377	T-cell antigen rec
39	19	23.2	13	2	S10562	zona pellucida-bin
40	19	23.2	13	2	JQ2309	hypothetical 1.6K
41	19	23.2	13	2	JQ2319	hypothetical 1.6K
42	19	23.2	14	2	B28018	very late antigen-
43	19	23.2	15	2	C48401	ribosomal protein
44	18	22.0	8	2	PT0030	inulinase (EC 3.2.
45	18	22.0	9	2	A61357	phyllocaerulein -

ALIGNMENTS

RESULT 1
S70332
endosperm protein, 10K - rye (fragment)
C;Species: Secale cereale (rye)
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S70332
R;Rocher, A.; Calero, M.; Soriano, F.; Mendez, E.
Biochim. Biophys. Acta 1295, 13-22, 1996
A;Title: Identification of major rye secalins as coeliac immunoreactive proteins.
A;Reference number: S70327; MUID:96283789; PMID:8679669
A;Accession: S70332
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-9 <ROC>

Query Match 32.9%; Score 27; DB 2; Length 9;
Best Local Similarity 66.7%; Pred. No. 2.8e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy         10 TGPYTF 15
               |||| :
Db          1 TGPYEY 6
```

RESULT 2
A56970
GLYMA1 - soybean (fragment)
C;Species: Glycine max (soybean)
C;Date: 02-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 09-Jul-2004
C;Accession: A56970
R;Codina, M.R.
submitted to the Protein Sequence Database, September 1995
A;Reference number: A56970
A;Accession: A56970
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <COD>
A;Cross-references: UNIPROT:Q7M285

Query Match 29.3%; Score 24; DB 2; Length 15;
Best Local Similarity 44.4%; Pred. No. 6.3e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy          5 LNNRYTGPY 13
              | : | ||:
Db          6 LADTYRGPF 14
```

RESULT 3
S08209
hypothetical protein 2 - garden pea
N;Alternate names: phytochrome

```
C;Species: Pisum sativum (garden pea)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Sep-1997
C;Accession: S08209
R;Sato, N.
Plant Mol. Biol. 11, 697-710, 1988
A;Title: Nucleotide sequence and expression of the phytochrome gene in Pisum sativum: di
A;Reference number: S06856
A;Accession: S08209
A;Molecule type: DNA
A;Residues: 1-15 <SAT>
A;Cross-references: EMBL:X14077; NID:g20836; PID:g20838
C;Genetics:
A;Gene: phy

  Query Match             28.0%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 9.5e+02;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 NRYTGPY 13
              | |  ||
Db          8 NGYFNPY 14


RESULT 4
A28856
fructose-bisphosphate aldolase (EC 4.1.2.13) B, hepatic - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 31-Oct-1997
C;Accession: A28856
R;Hannappel, E.; MacGregor, J.S.; Davoust, S.; Horecker, B.L.
Arch. Biochem. Biophys. 214, 293-298, 1982
A;Title: Limited proteolysis of liver and muscle aldolases: effects of subtilisin, cathe
A;Reference number: A28856; MUID:82205113; PMID:7044315
A;Accession: A28856
A;Molecule type: protein
A;Residues: 1-12 <HAN>
C;Superfamily: fructose-bisphosphate aldolase
C;Keywords: aldehyde-lyase; carbon-carbon lyase; gluconeogenesis; glycolysis; liver; per

  Query Match             26.8%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   42.9%;  Pred. No. 1.1e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 YTGPYTF 15
              :|  ||:
Db          6 FTASYTY 12


RESULT 5
PH1459
T-cell receptor beta chain (clone A3/IIC7) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Mar-2004
C;Accession: PH1459; PH1460
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Ko
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompatib
A;Reference number: PH1430; MUID:93171821; PMID:8436911
A;Accession: PH1459
A;Molecule type: mRNA
A;Residues: 1-12 <CA1>
A;Experimental source: cytolytic T-lymphocyte clone A3/IIC7
A;Accession: PH1460
A;Molecule type: mRNA
A;Residues: 1-12 <CA2>
A;Experimental source: cytolytic T-lymphocyte clone 332/2A
C;Keywords: receptor; T-cell

  Query Match             26.8%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          9 YTGPYTF 15
              |  |||
Db          5 YGSDYTF 11


RESULT 6
S36893
ribosomal protein - Mycobacterium bovis (fragment)
C;Species: Mycobacterium bovis
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C;Accession: S36893
R;Ohara, N.; Kimura, M.; Higashi, Y.; Yamada, T.
FEBS Lett. 331, 9-14, 1993
A;Title: Isolation and amino acid sequence of the 30S ribosomal protein S19 from Mycobact
A;Reference number: S36887; MUID:94009653; PMID:8405418
A;Accession: S36893
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <OHA>
A;Cross-references: UNIPROT:Q9R541

  Query Match             26.8%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 YTGPYT 14
              | ||| |
Db          3 YEGPKT 8


RESULT 7
A28497
neurotensin-related protein - turkey (fragment)
C;Species: Meleagris gallopavo (common turkey)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A28497
R;Carraway, R.E.; Cochrane, D.E.; Ruane, S.E.
J. Biol. Chem. 262, 15886-15889, 1987
A;Title: Isolation, structures, and biologic activity of neurotensin-related peptides ger
A;Reference number: A28497; MUID:88058942; PMID:2445741
A;Accession: A28497
A;Molecule type: protein
A;Residues: 1-15 <CAR>
A;Cross-references: UNIPROT:Q7LZA3
C;Keywords: neuropeptide

  Query Match             26.8%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 RYTGPY 13
              |  |||
Db          8 RTRGPY 13


RESULT 8
D47393
neuropeptide callatostatin 4 - bluebottle fly (Calliphora vomitoria)
C;Species: Calliphora vomitoria
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: D47393
R;Duve, H.; Johnsen, A.H.; Scott, A.G.; Yu, C.G.; Yagi, K.J.; Tobe, S.S.; Thorpe, A.
Proc. Natl. Acad. Sci. U.S.A. 90, 2456-2460, 1993
A;Title: Callatostatins: neuropeptides from the blowfly Calliphora vomitoria with sequenc
A;Reference number: A47393; MUID:93211980; PMID:8460157
A;Accession: D47393
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-8 <DUV>
A;Cross-references: UNIPROT:P41840
A;Experimental source: thoracic ganglia
A;Note: sequence extracted from NCBI backbone (NCBIP:128479)

  Query Match             25.6%;  Score 21;  DB 2;  Length 8;
```

```
  Best Local Similarity   75.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 PYTF 15
              ||:|
Db          3 PYSF 6


RESULT 9
PH1593
Ig H chain V-D-J region (wild-type clone 144) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1593
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609; PMID:8315387
A;Accession: PH1593
A;Molecule type: DNA
A;Residues: 1-13 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin

  Query Match             25.6%;  Score 21;  DB 2;  Length 13;
  Best Local Similarity   41.7%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 ARTLNNRYTGPY 13
              || | : |   |
Db          2 ARPLRHYYAMDY 13


RESULT 10
D28587
T-cell receptor beta-2 chain J-B2.5 segment - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 05-Nov-1999
C;Accession: D28587
R;Toyonaga, B.; Yoshikai, Y.; Vadasz, V.; Chin, B.; Mak, T.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 8624-8628, 1985
A;Title: Organization and sequences of the diversity, joining, and constant region genes
A;Reference number: A94081; MUID:86094276; PMID:3866244
A;Accession: D28587
A;Molecule type: DNA
A;Residues: 1-15 <TOY>
A;Cross-references: GB:M14159; NID:g338852; PIDN:AAA60679.1; PID:g553690
C;Keywords: T-cell receptor

  Query Match             25.6%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 RYTGPYT 14
              :| || |
Db          4 QYFGPGT 10


RESULT 11
F28587
T-cell receptor beta-2 chain J-B2.7 segment - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Aug-1988 #sequence_revision 16-Aug-1988 #text_change 05-Nov-1999
C;Accession: F28587
R;Toyonaga, B.; Yoshikai, Y.; Vadasz, V.; Chin, B.; Mak, T.W.
Proc. Natl. Acad. Sci. U.S.A. 82, 8624-8628, 1985
A;Title: Organization and sequences of the diversity, joining, and constant region genes
A;Reference number: A94081; MUID:86094276; PMID:3866244
A;Accession: F28587
A;Molecule type: DNA
A;Residues: 1-15 <TOY>
A;Cross-references: GB:M14159; NID:g338852; PIDN:AAA60681.1; PID:g553692
C;Keywords: T-cell receptor

  Query Match             25.6%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 RYTGPYT 14
              :| || |
Db          4 QYFGPGT 10


RESULT 12
I53284
T-cell receptor beta 2 chain J region, Jbeta2.7 - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C;Accession: I53284
R;Harindranath, N.; Alexander, C.B.; Mage, R.G.
Mol. Immunol. 28, 881-888, 1991
A;Title: Evolutionarily conserved organization and sequences of germline diversity and jo
A;Reference number: A53284; MUID:91342695; PMID:1678859
A;Accession: I53284
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-15 <HAR>
A;Cross-references: GB:S60737; NID:g233916; PIDN:AAB19525.1; PID:g233925
A;Note: sequence extracted from NCBI backbone (NCBIN:60737, NCBIP:60747)
C;Keywords: T-cell receptor

  Query Match             25.6%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 RYTGPYT 14
              :| || |
Db          4 QYFGPGT 10


RESULT 13
I38335
hypothetical TEL/MN1 mutant fusion protein type II - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Apr-2000
C;Accession: I38335
R;Buijs, A.; Sherr, S.; van Baal, S.; van Bezouw, S.; van der Plas, D.; Van Kessel, A.G.,
Oncogene 10, 1511-1519, 1995
A;Title: Translocation (12;22) (p13;q11) in myeloproliferative disorders results in fusio
A;Reference number: I38031; MUID:95249265; PMID:7731705
A;Accession: I38335
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-15 <BUI>
A;Cross-references: EMBL:X85024; NID:g971471; PIDN:CAA59397.1; PID:g971472
C;Comment: This sequence is the chimeric product of a translocation mutation.
C;Genetics:
A;Gene: ETV6/MN1; TEL/MN1
A;Map position: 22q11/12p13
C;Keywords: fusion protein

  Query Match             25.6%;  Score 21;  DB 4;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 2.1e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 RYTGPYT 14
              ||   |::
Db          1 RYRSPHS 7


RESULT 14
PH1462
T-cell receptor beta chain (clone A24/PEF1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Mar-2004
C;Accession: PH1462
```

R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Kd
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompatib
A;Reference number: PH1430; MUID:93171821; PMID:8436911
A;Accession: PH1462
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Keywords: receptor; T-cell

Query Match 25.0%; Score 20.5; DB 2; Length 12;
Best Local Similarity 83.3%; Pred. No. 2e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```
Qy         11 GP-YTF 15
                || |||
Db          6 GPDYTF 11
```

RESULT 15
PH0799
T-cell receptor alpha chain (H1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0799
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0799
A;Molecule type: mRNA
A;Residues: 1-13 <CAS>
A;Cross-references: EMBL:X60905
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 25.0%; Score 20.5; DB 2; Length 13;
Best Local Similarity 50.0%; Pred. No. 2.2e+03;
Matches 7; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

```
Qy          2 ARTLNNRYTGPYTF 15
              ||  |   ||  ||
Db          2 ARGTN---TGKLTF 12
```

Search completed: February 22, 2005, 09:46:26
Job time : 12.1333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-6
Perfect score: 82
Sequence: 1 SARTLNNRYTGPYTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	32.9	8	1	AL15_CARMA	P81818 carcinus ma
2	26	31.7	8	1	ALL8_CARMA	P81811 carcinus ma
3	26	31.7	8	1	ALL9_CARMA	P81812 carcinus ma
4	26	31.7	12	2	Q41856	Q41856 zea mays (m
5	25	30.5	10	2	Q6EX62	Q6ex62 hyptis suav
6	25	30.5	13	2	P90442	P90442 spodoptera
7	25	30.5	15	2	Q9QV00	Q9qv00 rattus sp.
8	24	29.3	8	1	AL16_CARMA	P81819 carcinus ma
9	24	29.3	15	2	Q9UR63	Q9ur63 emericella
10	24	29.3	15	2	Q7M285	Q7m285 glycine max
11	23.5	28.7	15	2	P97249	P97249 nitrosospir
12	23	28.0	9	1	AL11_CARMA	P81814 carcinus ma
13	23	28.0	15	2	Q9R4M8	Q9r4m8 bacillus fi
14	23	28.0	15	2	Q924T1	Q924t1 rattus norv
15	22	26.8	10	1	CA12_LITCI	P62540 litoria cit
16	22	26.8	10	1	CA12_LITSP	P62541 litoria spl
17	22	26.8	14	2	Q9MRF6	Q9mrf6 hordeum mur
18	22	26.8	15	2	Q9R541	Q9r541 mycobacteri
19	22	26.8	15	2	Q7LZA3	Q7lza3 meleagris g
20	21	25.6	8	1	ALL3_CYDPO	P82154 cydia pomon
21	21	25.6	8	1	ALL4_CALVO	P41840 calliphora
22	21	25.6	8	1	ALL4_CYDPO	P82155 cydia pomon
23	21	25.6	9	1	FLA2_TREHY	P80159 treponema h
24	21	25.6	10	2	Q6A3T8	Q6a3t8 archangiopt
25	21	25.6	10	2	Q71V02	Q71v02 pseudomonas
26	21	25.6	13	1	MP1_MICOC	P81532 microplitis
27	21	25.6	15	2	Q95751	Q95751 brachylophu
28	21	25.6	15	2	Q9R564	Q9r564 escherichia
29	20	24.4	7	1	ALL3_CARMA	P81806 carcinus ma
30	20	24.4	7	1	ALL4_CARMA	P81807 carcinus ma
31	20	24.4	7	1	ALL5_CARMA	P81808 carcinus ma
32	20	24.4	7	1	MNP1_LEPDE	P42984 leptinotars
33	20	24.4	8	1	AL17_CARMA	P81820 carcinus ma
34	20	24.4	8	1	ALL7_CARMA	P81809 carcinus ma
35	20	24.4	9	1	AL10_CARMA	P81813 carcinus ma
36	20	24.4	11	1	TKC2_CALVO	P41518 calliphora
37	20	24.4	12	2	Q7RW16	Q7rw16 neurospora
38	20	24.4	13	2	Q7M2Z6	Q7m2z6 ovis aries
39	20	24.4	13	2	Q6BCX4	Q6bcx4 oncorhynchu
40	20	24.4	14	2	Q8MH00	Q8mh00 homo sapien
41	20	24.4	14	2	Q8MH01	Q8mh01 homo sapien
42	20	24.4	14	2	Q8MH02	Q8mh02 homo sapien
43	20	24.4	14	2	Q8MH03	Q8mh03 homo sapien
44	20	24.4	14	2	Q8MH04	Q8mh04 homo sapien
45	20	24.4	14	2	Q8MH05	Q8mh05 homo sapien

ALIGNMENTS

RESULT 1
AL15_CARMA
ID AL15_CARMA STANDARD; PRT; 8 AA.
AC P81818;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Carcinustatin 15.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC -!- SIMILARITY: Belongs to the allatostatin family.
KW Amidation; Direct protein sequencing; Multigene family; Neuropeptide.
FT MOD_RES 8 8 Leucine amide.
SQ SEQUENCE 8 AA; 811 MW; 922879D5AB47687D CRC64;

Query Match 32.9%; Score 27; DB 1; Length 8;
Best Local Similarity 80.0%; Pred. No. 1.6e+06;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy         11 GPYTF 15
              |||:|
Db          2 GPYSF 6
```

RESULT 2
ALL8_CARMA
ID ALL8_CARMA STANDARD; PRT; 8 AA.
AC P81811;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Carcinustatin 8.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX NCBI_TaxID=6759;
RN [1]
RP SEQUENCE.
RC TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX MEDLINE=98121193; PubMed=9461295;

```
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Amidation; Direct protein sequencing; Multigene family; Neuropeptide.
FT   MOD_RES       8       8       Leucine amide.
SQ   SEQUENCE   8 AA;  795 MW;  922879CDCB47687D CRC64;

  Query Match             31.7%;  Score 26;  DB 1;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 GPYTF 15
              |||| |
Db          2 GPYAF 6


RESULT 3
ALL9_CARMA
ID   ALL9_CARMA     STANDARD;      PRT;     8 AA.
AC   P81812;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Carcinustatin 9.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Amidation; Direct protein sequencing; Multigene family; Neuropeptide.
FT   MOD_RES       8       8       Leucine amide.
SQ   SEQUENCE   8 AA;  781 MW;  7C2879CDCB476878 CRC64;

  Query Match             31.7%;  Score 26;  DB 1;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 GPYTF 15
              |||| |
Db          2 GPYAF 6


RESULT 4
Q41856
ID   Q41856      PRELIMINARY;      PRT;    12 AA.
AC   Q41856;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Sucrose synthetase gene mutant (Sh-5586 allele) containing transposon
DE   Tz86, 5' end of Tz86. (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85153004; PubMed=6099243;
RA   Dellaporta S.L., Chomet P.S., Mottinger J.P., Wood J.A., Yu S.M.,
RA   Hicks J.B.;
RT   "Endogenous transposable elements associated with virus infection in
RT   maize.";
RL   Cold Spring Harb. Symp. Quant. Biol. 49:321-328(1984).
DR   EMBL; M10174; AAA33516.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   12 AA;  1239 MW;  925ECBFEB0A5B861 CRC64;

  Query Match             31.7%;  Score 26;  DB 2;  Length 12;
  Best Local Similarity   62.5%;  Pred. No. 9.9e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 NRYTGPYT 14
              :|||| |
Db          2 SRYTGSST 9


RESULT 5
Q6EX62
ID   Q6EX62      PRELIMINARY;      PRT;    10 AA.
AC   Q6EX62;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Ribosomal protein (Fragment).
GN   Name=rps16;
OS   Hyptis suaveolens.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Hyptis.
OX   NCBI_TaxID=204129;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Willis F.,
RA   Powell M.P., Savolainen V.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ505341; CAD45464.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Chloroplast; Ribosomal protein.
FT   NON_TER       1       1
FT   NON_TER      10      10
SQ   SEQUENCE   10 AA;  1087 MW;  3783107729D1BB47 CRC64;

  Query Match             30.5%;  Score 25;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 PYTF 15
              ||||
Db          3 PYTF 6


RESULT 6
P90442
ID   P90442      PRELIMINARY;      PRT;    13 AA.
AC   P90442;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Polyhedrin (Fragment).
OS   Spodoptera littoralis nuclear polyhedrosis virus (SlNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=10456;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97300849; PubMed=9155869;
RA   Faktor O., Toister-Achituv M., Nachum O.;
RT   "Enhancer element, repetitive sequences and gene organization in an 8-
```

```
RT   kbp region containing the polyhedrin gene of the Spodoptera littoralis
RT   nucleopolyhedrovirus.";
RL   Arch. Virol. 142:1-15(1997).
DR   EMBL; X99711; CAA68046.1; -.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR001746; Polyhedrin.
DR   Pfam; PF00738; Polyhedrin; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   13 AA;  1383 MW;  2B0CD62832655737 CRC64;

  Query Match             30.5%;  Score 25;  DB 2;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 YTGP 12
              ||||
Db          8 YTGP 11


RESULT 7
Q9QV00
ID   Q9QV00      PRELIMINARY;      PRT;    15 AA.
AC   Q9QV00;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Eosinophil cationic protein (Fragment).
OS   Rattus sp. .
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10118;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95375557; PubMed=7647580;
RA   Watanabe M., Nittoh T., Suzuki T., Kitoh A., Mue S., Ohuchi K.;
RT   "Isolation and partial characterization of eosinophil granule proteins
RT   in rats--eosinophil cationic protein and major basic protein.";
RL   Int. Arch. Allergy Immunol. 108:11-18(1995).
SQ   SEQUENCE   15 AA;  1785 MW;  2AF6D99C12EBD794 CRC64;

  Query Match             30.5%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 NRYTG 11
              ||:||
Db          4 NRFTG 8


RESULT 8
AL16_CARMA
ID   AL16_CARMA     STANDARD;      PRT;     8 AA.
AC   P81819;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Carcinustatin 16.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Amidation; Direct protein sequencing; Multigene family; Neuropeptide.
FT   MOD_RES       8       8       Leucine amide.
SQ   SEQUENCE   8 AA;  813 MW;  7C286B45AB476878 CRC64;

  Query Match             29.3%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+06;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 GPYTF 15
              |||::
Db          2 GPYSY 6


RESULT 9
Q9UR63
ID   Q9UR63      PRELIMINARY;      PRT;    15 AA.
AC   Q9UR63;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Beta-D-FRUCTOFURANOSIDE FRUCTOHYDROLASE 60 kDa high molecular weight
DE   isoform (EC 3.2.1.26) (Fragment).
OS   Emericella nidulans (Aspergillus nidulans).
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC   Eurotiales; Trichocomaceae; Emericella.
OX   NCBI_TaxID=162425;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96409246; PubMed=8814228; DOI=10.1016/0167-4838(96)00073-8;
RA   Chen J.S., Saxton J., Hemming F.W., Peberdy J.F.;
RT   "Purification and partial characterization of the high and low
RT   molecular weight form (S- and F-form) of invertase secreted by
RT   Aspergillus nidulans.";
RL   Biochim. Biophys. Acta 1296:207-218(1996).
DR   GO; GO:0004564; F:beta-fructofuranosidase activity; IEA.
SQ   SEQUENCE   15 AA;  1588 MW;  2C992B42211366BB CRC64;

  Query Match             29.3%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 YTGPY 13
              || ||
Db         11 YTEPY 15


RESULT 10
Q7M285
ID   Q7M285      PRELIMINARY;      PRT;    15 AA.
AC   Q7M285;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   GLYMA1 (Fragment).
OS   Glycine max (Soybean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX   NCBI_TaxID=3847;
RN   [1]
RP   SEQUENCE.
RA   Codina M.R.;
RL   Submitted (SEP-1995) to the PIR data bank.
DR   PIR; A56970; A56970.
FT   NON_TER       1       1
FT   NON_TER      15      15
SQ   SEQUENCE   15 AA;  1795 MW;  9F02F7DE3A760BD7 CRC64;

  Query Match             29.3%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 LNNRYTGPY 13
```

```
              | : | ||:
Db          6 LADTYRGPF 14


RESULT 11
P97249
ID   P97249      PRELIMINARY;      PRT;    15 AA.
AC   P97249;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Ammonia monooxygenase subunit AmoC3 (Ammonia monooxygenase subunit
DE   AmoC2) (Fragment).
GN   Name=amoC3; Synonyms=amoC2;
OS   Nitrosospira multiformis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC   Nitrosomonadaceae; Nitrosospira.
OX   NCBI_TaxID=1231;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 25196;
RX   MEDLINE=21665709; PubMed=11807563;
RA   Norton J.M., Alzerreca J.J., Suwa Y., Klotz M.G.;
RT   "Diversity of ammonia monooxygenase operon in autotrophic ammonia-
RT   oxidizing bacteria.";
RL   Arch. Microbiol. 177:139-149(2002).
DR   EMBL; U89833; AAB48532.1; -.
DR   EMBL; U15733; AAB48013.1; -.
DR   GO; GO:0004497; F:monooxygenase activity; IEA.
KW   Monooxygenase.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1790 MW;  030B79929D28A467 CRC64;

  Query Match             28.7%;  Score 23.5;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+03;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          5 LNNRYTGPY 13
              ||||   ||
Db          4 LNNRIV-PY 11


RESULT 12
AL11_CARMA
ID   AL11_CARMA     STANDARD;      PRT;     9 AA.
AC   P81814;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Carcinustatin 11.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: May act as a neurotransmitter or neuromodulator.
CC   -!- SIMILARITY: Belongs to the allatostatin family.
KW   Amidation; Direct protein sequencing; Multigene family; Neuropeptide.
FT   MOD_RES       9      9       Leucine amide.
SQ   SEQUENCE   9 AA;  927 MW;  832D79CDCB46D861 CRC64;

  Query Match             28.0%;  Score 23;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         10 TGPYTF 15
              || | |
Db          2 TGQYAF 7


RESULT 13
Q9R4M8
ID   Q9R4M8      PRELIMINARY;      PRT;    15 AA.
AC   Q9R4M8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   Catalase isozyme I (Fragment).
OS   Bacillus firmus.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1399;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95267795; PubMed=7748885; DOI=10.1016/0005-2728(95)00016-C;
RA   Hicks D.B.;
RT   "Purification of three catalase isozymes from facultatively
RT   alkaliphilic Bacillus firmus OF4.";
RL   Biochim. Biophys. Acta 1229:347-355(1995).
SQ   SEQUENCE   15 AA;  1677 MW;  12E47DC8F66876ED CRC64;

  Query Match             28.0%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   46.2%;  Pred. No. 4.4e+03;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    2;  Gaps    1;

Qy          4 TLNNRYTG--PYT 14
              | ||   |  |:|
Db          3 TQNNENAGKXPFT 15


RESULT 14
Q924T1
ID   Q924T1      PRELIMINARY;      PRT;    15 AA.
AC   Q924T1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DNase I (Fragment).
GN   Name=DNASE1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21226133; PubMed=11327695; DOI=10.1006/bbrc.2001.4770;
RA   Yasuda T., Takeshita H., Ueki M., Iida R., Nakajima T., Mori S.,
RA   Mogi K., Kaneko Y., Kishi K.;
RT   "Tissue-specific in vivo inhibition of DNase I gene expression by
RT   somatostatin.";
RL   Biochem. Biophys. Res. Commun. 283:287-291(2001).
DR   EMBL; AB057442; BAB62091.1; -.
FT   NON_TER       15     15
SQ   SEQUENCE   15 AA;  1743 MW;  35D288BAED1DE561 CRC64;

  Query Match             28.0%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 4.4e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 RYTG 11
              ||||
Db          2 RYTG 5


RESULT 15
CA12_LITCI
ID   CA12_LITCI     STANDARD;      PRT;    10 AA.
AC   P62540; P82086;
```

```
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Caerulein 1.2/1.2Y4.
OS   Litoria citropa (Australian blue mountains tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=94770;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RX   MEDLINE=20057701; PubMed=10589099;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J.;
RT   "Caerulein-like peptides from the skin glands of the Australian blue
RT   mountains tree frog Litoria citropa. Part 1. Sequence determination
RT   using electrospray mass spectrometry.";
RL   Rapid Commun. Mass Spectrom. 13:2498-2502(1999).
CC   -!- FUNCTION: Hypotensive neuropeptide (Probable).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Skin dorsal glands.
CC   -!- PTM: Isoform 1.2Y4 differs from isoform 1.2 in not being sulfated.
CC   -!- MASS SPECTROMETRY: MW=1366; METHOD=Electrospray; RANGE=1-10
CC       (Isoform caerulein 1.2); NOTE=Ref.1.
CC   -!- MASS SPECTROMETRY: MW=1286; METHOD=Electrospray; RANGE=1-10
CC       (Isoform caerulein 1.2Y4); NOTE=Ref.1.
CC   -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
DR   InterPro; IPR001651; Gastrin.
DR   PROSITE; PS00259; GASTRIN; FALSE_NEG.
KW   Amidation; Amphibian defense peptide; Direct protein sequencing;
KW   Hypotensive agent; Pyrrolidone carboxylic acid; Sulfation.
FT   MOD_RES       1      1       Pyrrolidone carboxylic acid.
FT   MOD_RES       4      4       Sulfotyrosine (in form caerulein 1.2).
FT   MOD_RES      10     10       Phenylalanine amide.
SQ   SEQUENCE   10 AA;  1306 MW;  99DBFCD37861BB5A CRC64;

  Query Match             26.8%;  Score 22;  DB 1;  Length 10;
  Best Local Similarity   57.1%;  Pred. No. 4.2e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 YTGPYTF 15
              ||| : |
Db          4 YTGWFDF 10


Search completed: February 22, 2005, 09:37:54
Job time : 53.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-6
Perfect score: 82
Sequence: 1 SARTLNNRYTGPYTF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	82	100.0	15	2	AAW04846	Aaw04846 Self epit
2	36	43.9	10	3	AAY77563	Aay77563 GPCR-B4 p
3	34	41.5	13	5	AAE20719	Aae20719 Human Mls
4	34	41.5	13	5	AAE21020	Aae21020 Human Icr
5	31	37.8	12	2	AAW23267	Aaw23267 Tumour ne
6	31	37.8	12	2	AAW23269	Aaw23269 Tumour ne
7	31	37.8	13	2	AAW23268	Aaw23268 Tumour ne
8	31	37.8	13	8	ADP44318	Adp44318 Chaperone
9	31	37.8	14	6	ABR55893	Abr55893 CDR3-junc
10	31	37.8	15	6	ABR33651	Abr33651 Human can
11	31	37.8	15	6	ABR33761	Abr33761 Human can
12	31	37.8	15	6	ABR33743	Abr33743 Human can
13	30	36.6	12	2	AAW16032	Aaw16032 Peptide c
14	30	36.6	12	2	AAW68670	Aaw68670 Peptide b
15	30	36.6	12	2	AAW58149	Aaw58149 IL-1RtI b
16	30	36.6	12	2	AAY09793	Aay09793 Interleuk
17	30	36.6	12	3	AAB17593	Aab17593 IL-1 anta
18	30	36.6	12	5	ABB72489	Abb72489 Interleuk
19	30	36.6	12	7	ADJ72682	Adj72682 Interleuk
20	30	36.6	12	8	ADJ52317	Adj52317 CH1 delet
21	30	36.6	12	8	ADJ51280	Adj51280 CH1 delet
22	30	36.6	13	7	ADM75241	Adm75241 Potential
23	29	35.4	10	2	AAR67679	Aar67679 Go-alpha-
24	29	35.4	10	3	AAY79428	Aay79428 Tie2 rece
25	29	35.4	10	7	ADM08993	Adm08993 Canine im
26	29	35.4	10	7	ADM08999	Adm08999 Canine im
27	29	35.4	12	3	AAY56630	Aay56630 Virus-lik
28	29	35.4	12	3	AAY56273	Aay56273 Human cat
29	29	35.4	13	3	AAB39640	Aab39640 Anti-IL12
30	29	35.4	14	7	ADC83481	Adc83481 Asparagin
31	29	35.4	14	7	ADC83531	Adc83531 Asparagin
32	29	35.4	14	7	ADC83579	Adc83579 Asparagin
33	28	34.1	9	6	ABO10713	Abo10713 Murine J4
34	28	34.1	9	6	ABR44657	Abr44657 Murine J4
35	28	34.1	9	8	ADQ90785	Adq90785 Mouse com
36	28	34.1	10	4	AAU25785	Aau25785 Breast ca
37	28	34.1	11	8	ADE25847	Ade25847 Anti-alph
38	28	34.1	14	2	AAR60752	Aar60752 Lolium pe
39	28	34.1	14	7	ADC83623	Adc83623 Asparagin
40	28	34.1	14	8	ADH89412	Adh89412 Human tra
41	27	32.9	6	2	AAR25089	Aar25089 bGRF prod
42	27	32.9	8	2	AAR25090	Aar25090 bGRF prod
43	27	32.9	9	2	AAW27420	Aaw27420 CDR3 from
44	27	32.9	9	4	AAG63652	Aag63652 Complemen
45	27	32.9	9	6	ABJ38610	Abj38610 Hepatitis

ALIGNMENTS

RESULT 1
AAW04846
ID AAW04846 standard; peptide; 15 AA.
XX
AC AAW04846;
XX
DT 18-FEB-1997 (first entry)
XX
DE Self epitope of desmoglein 3, implicated in autoimmune disease.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Homo sapiens.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 1; Page 41; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is derived from the human desmoglein

```
CC   3 protein (amino acids 512-526) and is implicated as a self epitope in
CC   pemphigus vulgaris. Peptides derived from the human desmoglein protein
CC   are described in AAW04841-47
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 82;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-07;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 SARTLNNRYTGPYTF 15
              |||||||||||||||
Db          1 SARTLNNRYTGPYTF 15


RESULT 2
AAY77563
ID   AAY77563 standard; peptide; 10 AA.
XX
AC   AAY77563;
XX
DT   08-MAY-2000  (first entry)
XX
DE   GPCR-B4 peptide fragment.
XX
KW   Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; rat;
KW   taste transduction pathway; taste receptor; foliate; fungiform; food;
KW   circumvallate; taste signaling; pharmaceutical.
XX
OS   Rattus sp.
OS   Mus sp.
XX
PN   WO200006593-A1.
XX
PD   10-FEB-2000.
XX
PF   27-JUL-1999;   99WO-US017104.
XX
PR   28-JUL-1998;   98US-0095464P.
PR   17-DEC-1998;   98US-0112747P.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Zuker CS,  Adler JE,  Lindemeier J;
XX
DR   WPI; 2000-195257/17.
XX
PT   New isolated sensory transduction G-protein coupled receptor, useful for
PT   developing products for use in studying and modulating the taste
PT   transduction pathway and for generating taste topographic maps.
XX
PS   Claim 7; Page 58; 76pp; English.
XX
CC   The invention provides nucleic acids encoding rat, mouse and human
CC   sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
CC   GPCR polypeptides are components of the taste transduction pathway. The
CC   nucleic acids can be used to identify taste cells and as tools for the
CC   generation of taste topographic maps that elucidate the relationship
CC   between the taste cells of the tongue and taste sensory neurons leading
CC   to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
CC   for identifying subpopulations of taste receptor cells such as foliate,
CC   fungiform, and circumvallate taste receptor cells. The polypeptides can
CC   be used for identifying compounds that modulate sensory signaling in
CC   sensory cells. Such modulators of taste transduction are useful for
CC   pharmacological and genetic modulation of taste signaling pathways. These
CC   modulatory compounds can then be used in the food and pharmaceutical
CC   industries to customize taste. The present sequence represents a GPCR-B4
CC   peptide fragment
XX
SQ   Sequence 10 AA;

  Query Match             43.9%;  Score 36;  DB 3;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 17;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 RYTGPYTF 15
              || ||| |
Db          3 RYHGPYVF 10


RESULT 3
AAE20719
ID   AAE20719 standard; peptide; 13 AA.
XX
AC   AAE20719;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Human Mlsn1 intra-cellular loop 2 peptide #1.
XX
KW   Human; Mlsn1 protein; calcium channel; immune system; macrophage; T-cell;
KW   monocyte; B-cell; mast cell; lcrac; immuno degenerative disorder; COPD;
KW   biomarker; autoimmune disease; inflammatory disease; organ rejection;
KW   arthritis; antisense gene therapy; intra-cellular loop 2.
XX
OS   Homo sapiens.
XX
PN   EP1184457-A1.
XX
PD   06-MAR-2002.
XX
PF   03-SEP-2001; 2001EP-00402283.
XX
PR   05-SEP-2000; 2000EP-00402436.
XX
PA   (WARN ) WARNER LAMBERT CO.
XX
PI   Schindler V,  Bloes C,  Fink M,  Allen J,  Grentzmann G;
XX
DR   WPI; 2002-271038/32.
DR   N-PSDB; AAD33208.
XX
PT   Novel Mlsn1 polypeptide and nucleic acid encoding the polypeptide, useful
PT   for screening Mlsn1 modulators and for diagnosing disorders associated
PT   with aberrant lcrac function in immune cells.
XX
PS   Claim 16; Page 37; 63pp; English.
XX
CC   The patent discloses novel human Mlsn1 polypeptides and polynucleotides
CC   encoding them. The invention also relates to the characterisation of
CC   capacitative calcium channel homologue for the activation of immune
CC   system. Mlsn1 gene products contain a novel calcium pore sequence which
CC   is a part of the capacitative calcium channel in immune cells such as
CC   macrophages, monocytes, T-cells, B-cells and mast cells. Mlsn1 sequences
CC   are useful for screening Mlsn1 modulators and for diagnosing disorders
CC   associated with aberrant lcrac function in immune cells. They are useful
CC   for screening ligand substances or molecules that are able to modulate
CC   the biological activity of Mlsn1. The compound identified as modulator is
CC   useful for the preparation of pharmaceutical compounds or compositions
CC   that bind and/or modulate the biological activity of Mlsn1. Sequences of
CC   the invention are useful for treating immuno degenerative disorders. They
CC   are also useful for developing biomarkers for immune system activation or
CC   autoimmune or inflammatory diseases (e.g., arthritis, organ rejection,
CC   COPD), as modulators of immune system activation and as therapeutic
CC   targets. Mlsn1 sequences are also used for antisense gene therapy. The
CC   present sequence is human mlsn1 intra-cellular loop 2 peptide
XX
SQ   Sequence 13 AA;

  Query Match             41.5%;  Score 34;  DB 5;  Length 13;
  Best Local Similarity   71.4%;  Pred. No. 53;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 NRYTGPY 13
              |:| |||
Db          1 NKYLGPY 7
```

```
RESULT 4
AAE21020
ID   AAE21020 standard; peptide; 13 AA.
XX
AC   AAE21020;
XX
DT   01-JUL-2002  (first entry)
XX
DE   Human Icrac-Mlsnl protein fragment #6.
XX
KW   Human; Icrac-Mlsnl; major capacitative calcium channel; ligand screening;
KW   macrophage maturation; detection.
XX
OS   Homo sapiens.
XX
PN   EP1186661-A1.
XX
PD   13-MAR-2002.
XX
PF   05-SEP-2000; 2000EP-00402436.
XX
PR   05-SEP-2000; 2000EP-00402436.
XX
PA   (WARN ) WARNER LAMBERT CO.
XX
PI   Allen J,  Grentzmann G,  Fink M,  Schindler V,  Bloes C;
XX
DR   WPI; 2002-282883/33.
DR   N-PSDB; AAD33455.
XX
PT   Polypeptide regions of major capacitative calcium channel and
PT   polynucleotide regions useful for screening modulators of channel
PT   activity for regulating activity of antigen presenting cell especially of
PT   macrophages.
XX
PS   Claim 16; Page 36; 62pp; English.
XX
CC   The invention relates to a purified or isolated N-terminal polypeptide
CC   region, C-terminal polypeptide region, extracellular and intracellular
CC   loop regions, of human Icrac-Mlsnl (major capacitative calcium channel).
CC   Icrac-Mlsnl DNA and protein are useful for screening ligand substances or
CC   molecules that are able to bind to a Icrac-Mlsnl and modulates the
CC   activity of antigen presenting cells especially macrophages maturation.
CC   Icrac-Mlsnl is also useful for screening ligand substances or molecules
CC   that are able to modulate the biological activity of Icrac-Mlsnl, where a
CC   change in calcium flux or activation potential within the exposed
CC   recombinant cell is measured, and where the recombinant cell expresses a
CC   reporter molecule under the control of Icrac-Mlsnl DNA, where a
CC   competitor Mlsnl binding ligand, is selected from the class of pyrazol
CC   compounds is added and the binding of the competitor in the presence
CC   and/or absence of a test compound is determined. Icrac-Mlsnl useful for
CC   detecting Icrac-Mlsnl DNA in a blood sample. The present sequence is
CC   human Icrac-Mlsnl protein intracellular loop 2
XX
SQ   Sequence 13 AA;

  Query Match             41.5%;  Score 34;  DB 5;  Length 13;
  Best Local Similarity   71.4%;  Pred. No. 53;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 NRYTGPY 13
              |:| |||
Db          1 NKYLGPY 7


RESULT 5
AAW23267
ID   AAW23267 standard; peptide; 12 AA.
XX
AC   AAW23267;
XX
DT   03-NOV-1997  (first entry)
XX
DE   Tumour necrosis factor alpha inhibiting peptide.
XX
KW   Tumour necrosis factor; alpha; TNF-alpha; inhibition; treatment;
KW   mediation; disorder; septic shock; bacterium; virus; fungus; infection;
KW   autoimmune; disease; alcohol induced hepatitis; sarcoiditis; Crohn's;
KW   disseminated intravascular coagulation; graft versus host; Rawasaki's;
KW   tumour; bacteria.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "acetylated"
FT   Modified-site   12
FT                   /note= "amidated"
XX
PN   US5641751-A.
XX
PD   24-JUN-1997.
XX
PF   01-MAY-1995;   95US-00432694.
XX
PR   01-MAY-1995;   95US-00432694.
XX
PA   (CENZ ) CENTOCOR INC.
XX
PI   Heavner GA;
XX
DR   WPI; 1997-340972/31.
XX
PT   Peptide inhibiting tumour necrosis factor alpha - useful for treating
PT   septic shock, infections, autoimmune diseases, etc.
XX
PS   Example 2; Col 23-24; 15pp; English.
XX
CC   The present peptide is a tumour necrosis factor alpha (TNF-alpha)
CC   inhibitor, which can be used to treat TNF-alpha mediated disorders, e.g.
CC   septic shock, bacterial, viral and fungal infections, autoimmune
CC   diseases, alcohol induced hepatitis, sarcoiditis, Crohn's disease,
CC   disseminated intravascular coagulation, graft versus host disease,
CC   Rawasaki's disease and TNF-alpha secreting tumours. The peptide is
CC   preferably given as a daily dose of 1-1000, preferably 1-10 mg/kg, and
CC   has an IC50 of 155 microM against binding of TNF-alpha to a p55TNFr-IgG
CC   construct
XX
SQ   Sequence 12 AA;

  Query Match             37.8%;  Score 31;  DB 2;  Length 12;
  Best Local Similarity   71.4%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LNNRYTG 11
              ::|||||
Db          6 VSNRYTG 12


RESULT 6
AAW23269
ID   AAW23269 standard; peptide; 12 AA.
XX
AC   AAW23269;
XX
DT   03-NOV-1997  (first entry)
XX
DE   Tumour necrosis factor alpha inhibiting peptide.
XX
KW   Tumour necrosis factor; alpha; TNF-alpha; inhibition; treatment;
KW   mediation; disorder; septic shock; bacterium; virus; fungus; infection;
KW   autoimmune; disease; alcohol induced hepatitis; sarcoiditis; Crohn's;
KW   disseminated intravascular coagulation; graft versus host; Rawasaki's;
KW   tumour; bacteria.
```

```
XX
OS   Synthetic.
XX
PN   US5641751-A.
XX
PD   24-JUN-1997.
XX
PF   01-MAY-1995;   95US-00432694.
XX
PR   01-MAY-1995;   95US-00432694.
XX
PA   (CENZ ) CENTOCOR INC.
XX
PI   Heavner GA;
XX
DR   WPI; 1997-340972/31.
XX
PT   Peptide inhibiting tumour necrosis factor alpha - useful for treating
PT   septic shock, infections, autoimmune diseases, etc.
XX
PS   Disclosure; Col 23-24; 15pp; English.
XX
CC   The present peptide is a tumour necrosis factor alpha (TNF-alpha)
CC   inhibitor, which can be used to treat TNF-alpha mediated disorders, e.g.
CC   septic shock, bacterial, viral and fungal infections, autoimmune
CC   diseases, alcohol induced hepatitis, sarcoiditis, Crohn's disease,
CC   disseminated intravascular coagulation, graft versus host disease,
CC   Rawasaki's disease and TNF-alpha secreting tumours. The peptide is
CC   preferably given as a daily dose of 1-1000, preferably 1-10 mg/kg
XX
SQ   Sequence 12 AA;

  Query Match             37.8%;  Score 31;  DB 2;  Length 12;
  Best Local Similarity   71.4%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LNNRYTG 11
              ::|||||
Db          6 VSNRYTG 12


RESULT 7
AAW23268
ID   AAW23268 standard; peptide; 13 AA.
XX
AC   AAW23268;
XX
DT   03-NOV-1997  (first entry)
XX
DE   Tumour necrosis factor alpha inhibiting peptide.
XX
KW   Tumour necrosis factor; alpha; TNF-alpha; inhibition; treatment;
KW   mediation; disorder; septic shock; bacterium; virus; fungus; infection;
KW   autoimmune; disease; alcohol induced hepatitis; sarcoiditis; Crohn's;
KW   disseminated intravascular coagulation; graft versus host; Rawasaki's;
KW   tumour; bacteria.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "acetylated"
FT   Modified-site   13
FT                   /note= "amidated"
XX
PN   US5641751-A.
XX
PD   24-JUN-1997.
XX
PF   01-MAY-1995;   95US-00432694.
XX
PR   01-MAY-1995;   95US-00432694.
XX
PA   (CENZ ) CENTOCOR INC.
XX
PI   Heavner GA;
XX
DR   WPI; 1997-340972/31.
XX
PT   Peptide inhibiting tumour necrosis factor alpha - useful for treating
PT   septic shock, infections, autoimmune diseases, etc.
XX
PS   Example 2; Col 23-24; 15pp; English.
XX
CC   The present peptide is a tumour necrosis factor alpha (TNF-alpha)
CC   inhibitor, which can be used to treat TNF-alpha mediated disorders, e.g.
CC   septic shock, bacterial, viral and fungal infections, autoimmune
CC   diseases, alcohol induced hepatitis, sarcoiditis, Crohn's disease,
CC   disseminated intravascular coagulation, graft versus host disease,
CC   Rawasaki's disease and TNF-alpha secreting tumours. The peptide is
CC   preferably given as a daily dose of 1-1000, preferably 1-10 mg/kg, and
CC   has an IC50 of 307 microM against binding of TNF-alpha to a p55TNFr-IgG
CC   construct
XX
SQ   Sequence 13 AA;

  Query Match             37.8%;  Score 31;  DB 2;  Length 13;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LNNRYTG 11
              ::|||||
Db          6 VSNRYTG 12


RESULT 8
ADP44318
ID   ADP44318 standard; peptide; 13 AA.
XX
AC   ADP44318;
XX
DT   26-AUG-2004  (first entry)
XX
DE   Chaperone-related peptide #5.
XX
KW   chaperone peptide; transmissible spongiform encephalopathy;
KW   Creutzfeld-Jakob disease; scrapie; bovine spongiform encephalopathy.
XX
OS   Synthetic.
XX
PN   JP2004155688-A.
XX
PD   03-JUN-2004.
XX
PF   05-NOV-2002; 2002JP-00321436.
XX
PR   30-APR-2002; 2002JP-00128976.
PR   10-JUL-2002; 2002JP-00200884.
PR   13-SEP-2002; 2002JP-00268260.
XX
PA   (BIOF-) BIO FRONTIER KENKYUSHO KK.
XX
DR   WPI; 2004-424357/40.
XX
PT   New synthetic peptide having chaperone activity, useful for treating
PT   transmissible spongiform encephalopathies such as Creutzfeld-Jakob
PT   disease and scrapie.
XX
PS   Example; SEQ ID NO 7; 22pp; Japanese.
XX
CC   The invention comprises synthetic peptides that have chaperone activity.
CC   The peptides of the invention are useful in the treatment of
CC   transmissible spongiform encephalopathies, Creutzfeld-Jakob disease,
CC   scrapie and bovine spongiform encephalopathy. The present amino acid
CC   sequence represents a peptide that was used in the exemplification of the
CC   invention.
```

```
XX
SQ   Sequence 13 AA;

  Query Match             37.8%;  Score 31;  DB 8;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 SARTLNNRY 9
              | || :|||
Db          5 SGRTWSNRY 13


RESULT 9
ABR55893
ID   ABR55893 standard; protein; 14 AA.
XX
AC   ABR55893;
XX
DT   02-SEP-2003  (first entry)
XX
DE   CDR3-junction sequence of kappa light chain of mAb 2C8.
XX
KW   Mitochondrial antigen; autoantibody; biliary cirrhosis; transgenic;
KW   monoclonal antibody; mAb.
XX
OS   Homo sapiens.
XX
PN   WO200276406-A2.
XX
PD   03-OCT-2002.
XX
PF   27-MAR-2002; 2002WO-US009694.
XX
PR   27-MAR-2001; 2001US-0279052P.
PR   21-SEP-2001; 2001US-0323920P.
XX
PA   (GERS/) GERSHWIN M E.
XX
PI   Gershwin ME;
XX
DR   WPI; 2003-018851/01.
XX
PT   New anti-mitochondrial antigen specific human monoclonal antibodies from
PT   patients with primary biliary cirrhosis (PBC), useful as a diagnostic
PT   reagent, or as a reagent for screening antagonists for treating patients
PT   with PBC.
XX
PS   Claim 34; Page 101; 150pp; English.
XX
CC   The invention relates to an isolated human antibody or its antigen-
CC   binding portions, which binds a mitochondrial antigen bound by a Human
CC   autoantibody found in patients with primary biliary cirrhosis. The
CC   antibody is useful as a diagnostic reagent, e.g. for monitoring whether a
CC   transplanted liver in a patient with primary biliary cirrhosis, is
CC   expressing a mitochondrial antigen bound by an autoantibody. The antibody
CC   is also useful for purifying an antigen to which it binds, for producing
CC   an anti-idiotype antibody to it, or as a reagent to screen for
CC   antagonists to it, either in vivo or in vitro. The antagonist is useful
CC   for treating a patient with primary biliary cirrhosis, or for
CC   prophylactically treating a patient about to have primary biliary
CC   cirrhosis. Sequences ABR55890-98 represent CDR3- junction sequences of
CC   the kappa light chains of the anti-mitochondrial antigen human monoclonal
CC   antibodies (mAbs) 2E6, 4D11, 1C7, 2C8, 3B4, 4G6, 4C2, 2E11 and 3B10
XX
SQ   Sequence 14 AA;

  Query Match             37.8%;  Score 31;  DB 6;  Length 14;
  Best Local Similarity   62.5%;  Pred. No. 2e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 RYTGPYTF 15
              ||  |:||
Db          4 RYNTPFTF 11


RESULT 10
ABR33651
ID   ABR33651 standard; peptide; 15 AA.
XX
AC   ABR33651;
XX
DT   19-MAY-2003  (first entry)
XX
DE   Human cancer-related protein 159P2B5 HLA peptide #41.
XX
KW   Human; cytostatic; vaccine; cancer; immune response; HLA;
KW   human leukocyte antigen.
XX
OS   Homo sapiens.
XX
PN   WO200283921-A2.
XX
PD   24-OCT-2002.
XX
PF   10-APR-2002; 2002WO-US011654.
XX
PR   10-APR-2001; 2001US-0282739P.
PR   10-APR-2001; 2001US-0283112P.
PR   25-APR-2001; 2001US-0286630P.
XX
PA   (AGEN-) AGENSYS INC.
XX
PI   Jakobovits A,  Challita-Eid PM,  Faris M,  Ge W,  Hubert RS;
PI   Morrison K,  Morrison RK,  Raitano AB;
XX
DR   WPI; 2003-075555/07.
XX
PT   New composition comprising a substance that modulates the structure of
PT   proteins and polynucleotides, useful for therapeutic, prognostic and
PT   diagnostic reagents for eliciting cellular or humoral immune response in
PT   cancer patients.
XX
PS   Claim 13; Page 533; 1021pp; English.
XX
CC   The present invention relates to novel human cancer-related genes and
CC   proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC   proteins are useful for eliciting a humoral or cellular immune response.
CC   The genes are useful as probes and primers for the amplification and/or
CC   detection of genes, mRNAs or their fragments, as reagents for the
CC   diagnosis and/or prognosis of cancer, as coding sequences capable of
CC   directing the expression of the protein, as tools for modulating or
CC   inhibiting the expression of genes and/or translation of transcripts, and
CC   as therapeutic agents. The proteins and peptides are useful as
CC   therapeutic, prognostic and diagnostic reagents for cancer. The present
CC   sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC   from the invention
XX
SQ   Sequence 15 AA;

  Query Match             37.8%;  Score 31;  DB 6;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RTLNNRYTGP 12
              ||   |:|||
Db          1 RTQARRHTGP 10


RESULT 11
ABR33761
ID   ABR33761 standard; peptide; 15 AA.
XX
AC   ABR33761;
XX
DT   19-MAY-2003  (first entry)
XX
```

DE Human cancer-related protein 159P2B5 HLA peptide #151.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 535; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 37.8%; Score 31; DB 6; Length 15;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 RTLNNRYTGP 12
|| |:|||
Db 3 RTQARRHTGP 12

RESULT 12
ABR33743
ID ABR33743 standard; peptide; 15 AA.
XX
AC ABR33743;
XX
DT 19-MAY-2003 (first entry)
XX
DE Human cancer-related protein 159P2B5 HLA peptide #133.
XX
KW Human; cytostatic; vaccine; cancer; immune response; HLA;
KW human leukocyte antigen.
XX
OS Homo sapiens.
XX
PN WO200283921-A2.
XX
PD 24-OCT-2002.
XX
PF 10-APR-2002; 2002WO-US011654.
XX
PR 10-APR-2001; 2001US-0282739P.
PR 10-APR-2001; 2001US-0283112P.
PR 25-APR-2001; 2001US-0286630P.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Challita-Eid PM, Faris M, Ge W, Hubert RS;
PI Morrison K, Morrison RK, Raitano AB;
XX
DR WPI; 2003-075555/07.
XX
PT New composition comprising a substance that modulates the structure of
PT proteins and polynucleotides, useful for therapeutic, prognostic and
PT diagnostic reagents for eliciting cellular or humoral immune response in
PT cancer patients.
XX
PS Claim 13; Page 535; 1021pp; English.
XX
CC The present invention relates to novel human cancer-related genes and
CC proteins (ABZ78120-ABZ78168 and ABR01789-ABR01861). The genes and
CC proteins are useful for eliciting a humoral or cellular immune response.
CC The genes are useful as probes and primers for the amplification and/or
CC detection of genes, mRNAs or their fragments, as reagents for the
CC diagnosis and/or prognosis of cancer, as coding sequences capable of
CC directing the expression of the protein, as tools for modulating or
CC inhibiting the expression of genes and/or translation of transcripts, and
CC as therapeutic agents. The proteins and peptides are useful as
CC therapeutic, prognostic and diagnostic reagents for cancer. The present
CC sequence is a human leukocyte antigen (HLA) peptide, used in an example
CC from the invention
XX
SQ Sequence 15 AA;

Query Match 37.8%; Score 31; DB 6; Length 15;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 RTLNNRYTGP 12
|| |:|||
Db 4 RTQARRHTGP 13

RESULT 13
AAW16032
ID AAW16032 standard; peptide; 12 AA.
XX
AC AAW16032;
XX
DT 19-AUG-1997 (first entry)
XX
DE Peptide containing QPY or QPY-like motif.
XX
KW interleukin-1; type I receptor; IL-1RtI; agonist; antagonist.
XX
OS Synthetic.
XX
PN WO9639165-A1.
XX
PD 12-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US009835.
XX
PR 05-JUN-1995; 95US-00464538.
XX
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Barrett RW, Yanofsky SD, Baldwin D, Jacobs JW, Bovy PR, Leahy EM;
PI Pottorf RS, Dharanipragada R, Tomlinson RC;
XX
DR WPI; 1997-042846/04.

XX
PT Interleukin-1 type I receptor inhibitor peptide(s) and compounds - used
PT to antagonise the activity of IL-1, for treatment of e.g. AIDS,
PT rheumatoid arthritis, chronic hepatitis B, etc.
XX
PS Disclosure; Page 22; 74pp; English.
XX
CC The invention relates to peptides which bind to the IL-1 type I receptor
CC and which comprise the motif WXXXG-Z1-W or the motif XXQ-Z5-Y-Z6-XX, in
CC which X can be any one of the 20 genetically coded L-amino acids or the
CC stereoisomeric D-amino acids or unnatural amino acids; Z1 is L, I, A or Q
CC ; Z5 is P or azetidine (presumably intended to be azetidine carboxylic
CC acid); and Z6 is S, A, V or L. The present sequence is that of a random
CC peptide from a library of peptides containing the "XXQ-Z5-Y-Z6-XX" motif,
CC the library being constructed to screen the peptides for activity. IL-1
CC type I receptor-binding peptides may be useful in the treatment of a
CC variety of IL-1 related disorders including atherosclerosis, rheumatoid
CC arthritis, osteoporosis, HIV infection and AIDS, bacterial infection,
CC respiratory distress syndrome, acute myelogenous leukaemia, graft versus
CC host disease, coal miner pneumoconiosis, alcoholic cirrhosis, cuprophane
CC haemodialysis, cardiopulmonary bypass, chronic hepatitis B, tuberculosis,
CC obstructive jaundice, Paget's disease and osteomalacia, IDDM, Kawasaki's
CC disease, inflammatory bowel disease, sepsis, toxic shock and luteal phase
XX
SQ Sequence 12 AA;

```
  Query Match             36.6%;  Score 30;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 NNRYTGPYTF 15
              |  |  ||:|
Db          1 NGNYWQPYSF 10
```

RESULT 14
AAW68670
ID AAW68670 standard; peptide; 12 AA.
XX
AC AAW68670;
XX
DT 01-OCT-1998 (first entry)
XX
DE Peptide binding interleukin-1 type I receptor.
XX
KW Antagonist; interleukin-1; IL-1; IL-1 type I receptor; IL-1RtI;
KW treatment; IL-1 disorder.
XX
OS Synthetic.
XX
PN US5786331-A.
XX
PD 28-JUL-1998.
XX
PF 05-JUN-1995; 95US-00465391.
XX
PR 02-FEB-1994; 94US-00190788.
PR 01-FEB-1995; 95US-00383474.
XX
PA (HMRI) HOECHST MARION ROUSSEL INC.
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Yanofsky SD, Jacobs JW, Bovy PR, Leahy EM, Barrett RW;
PI Pottorf RS, Baldwin D;
XX
DR WPI; 1998-436582/37.
XX
PT Antagonisation of action of interleukin-1 on type I receptor - by
PT contacting receptor with selected peptides.
XX
PS Disclosure; Col 13; 118pp; English.
XX
CC AAW68666-75 represent random peptide sequences, identified from a peptide
CC library based on the peptide AAW68899. The peptides antagonise the action
CC of interleukin-1 (IL-1) by binding to an IL-1 type I receptor (IL-1RtI).
CC The peptides are used in the treatment of disorders mediated by IL-1,
CC e.g. atherosclerosis, arthritis, osteoporosis, AIDS, bacterial
CC infections, respiratory distress syndrome, acute myelogenous leukaemia,
CC graft-versus-host disease, pneumoconiosis, cirrhosis, cuprophane
CC haemodialysis, cardiopulmonary bypass, hepatitis B, thermal injury,
CC reticulohistiocytosis, sarcoidosis, tuberculosis, obstructive jaundice,
CC Paget's disease, osteomalacia, diabetes, Kawasaki's disease, inflammatory
CC bowel disease, sepsis, toxic shock and luteal phase
XX
SQ Sequence 12 AA;

```
  Query Match             36.6%;  Score 30;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 NNRYTGPYTF 15
              |  |  ||:|
Db          1 NGNYWQPYSF 10
```

RESULT 15
AAW58149
ID AAW58149 standard; peptide; 12 AA.
XX
AC AAW58149;
XX
DT 14-AUG-1998 (first entry)
XX
DE IL-1RtI binding peptide SEQ ID NO:75 based in SEQ ID NO:45.
XX
KW Interleukin-1 type I receptor; binding peptide; IL-1R; IL-1RtI;
KW competitive inhibition; cytokine; blocker; IL-1 related disorder;
KW diagnosis; atherosclerosis; rheumatoid arthritis; osteoporosis.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN US5767234-A.
XX
PD 16-JUN-1998.
XX
PF 01-FEB-1995; 95US-00383474.
XX
PR 02-FEB-1994; 94US-00190788.
XX
PA (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI Baldwin D, Jacobs JW, Yanofsky SD, Barrett RW;
XX
DR WPI; 1998-361782/31.
XX
PT Peptide(s) that bind to interleukin-1 type I receptor - useful in
PT screening assays for interleukin receptors blockers, diagnosis and
PT therapy.
XX
PS Disclosure; Col 11; 89pp; English.
XX
CC A compound has been developed that binds to an interleukin-1 (IL-1) type
CC I receptor with an IC50 of 2.5 mM or less and has a molecular weight of
CC <3 kD, where the binding of the compound to the receptor is competitively
CC inhibited by a peptide of 8-25 amino acids comprising the core sequence:
CC Z7-Z8-Q-Z5-Y-Z6-Z9-Z10, where Z5 = Pro or azetidine; Z6 = Ser, Ala, Val
CC or Leu; Z7 = Tyr, Trp or Phe; Z8 = Glu, Phe, Val, Trp or Tyr; Z9 = Met,
CC Phe, Val, Arg, Gln, Lys, Thr, Ser, Asp, Leu, Ile or Glu; and Z10 = Glu,
CC Leu, Trp, Val, His, Ile, Gly, Ala, Asp, Tyr, Asn, Gln or Pro. Peptides of
CC the invention are used in screening assays for IL-1 receptor blockers.
CC They are also used as probes for detecting IL-1 type I receptor
CC expression on the surface of cells. The peptides are useful in treating
CC IL-1-related disorders, e.g. atherosclerosis, rheumatoid arthritis,
CC osteoporosis, HIV infection, AIDS, bacterial infection, respiratory
CC distress syndrome, acute myelogenous leukaemia (AML), graft versus host

CC disease, coal miner pneumonoconiosis, alcoholic cirrhosis, cuprophane
CC haemodialysis, cardiopulmonary bypass, chronic hepatitis B, thermal
CC injury, reticulohistiocytosis, sarcoidosis, tuberculosis, obstructive
CC jaundice, Paget's disease and osteomalacia, Kawasaki's disease,
CC inflammatory bowel disease, sepsis, toxic shock, and luteal phase. The
CC present sequence represents a peptide from the present invention
XX
SQ Sequence 12 AA;

Query Match 36.6%; Score 30; DB 2; Length 12;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy          6 NNRYTGPYTF 15
              |   |  ||:|
Db          1 NGNYWQPYSF 10
```

Search completed: February 22, 2005, 09:24:40
Job time : 68.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-7
Perfect score: 80
Sequence: 1 QSGTMRTRHSTGGTN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	35.0	11	2	A61512	variant surface gl
2	23	28.7	14	2	PH0776	T-cell receptor al
3	22	27.5	9	1	YFPG	thymic factor - pi
4	22	27.5	9	2	A60957	thymocyte growth p
5	22	27.5	15	2	PH0751	T-cell receptor be
6	21	26.2	8	2	PT0639	T-cell receptor be
7	21	26.2	11	2	PH0919	T-cell receptor be
8	21	26.2	11	2	S05002	corazonin - Americ
9	21	26.2	15	2	S57577	T cell receptor V-
10	20	25.0	8	2	A54823	olfactory receptor
11	20	25.0	12	2	D20907	Ig kappa-1 chain J
12	20	25.0	12	2	PH1464	T-cell receptor be
13	20	25.0	14	2	I51432	histone H4-1 precu
14	20	25.0	15	2	C43334	orf3 3' to aadR -
15	20	25.0	15	2	A28497	neurotensin-relate
16	19	23.8	6	2	PT0512	T-cell receptor be
17	19	23.8	9	2	PT0231	Ig heavy chain CDR
18	19	23.8	9	2	D44787	calliFMRFamide 13
19	19	23.8	10	2	S19296	16K protein - poul
20	19	23.8	10	2	E61512	variant surface gl
21	19	23.8	11	2	PT0287	Ig heavy chain CRD
22	19	23.8	12	2	A49033	T-cell receptor de
23	19	23.8	12	2	PH0931	T-cell receptor be
24	19	23.8	14	2	S58426	spermadhesin AWN h
25	19	23.8	15	2	A38304	heterogeneous ribo
26	19	23.8	15	2	PA0099	phenotipic variati
27	18	22.5	8	2	PT0588	T-cell receptor be
28	18	22.5	9	2	D28854	fibrinopeptide B -
29	18	22.5	10	2	B38887	T-cell receptor ga
30	18	22.5	11	2	PN0167	ribosomal protein
31	18	22.5	11	2	PT0302	Ig heavy chain CRD
32	18	22.5	11	2	B41946	T-cell receptor ga
33	18	22.5	11	2	C49037	TcR gamma V-J regi
34	18	22.5	12	2	C49033	T-cell receptor de
35	18	22.5	12	2	PT0216	T-cell receptor be
36	18	22.5	13	2	A23695	myosin heavy chain
37	18	22.5	13	2	S74130	NADH oxidase - Gia
38	18	22.5	14	2	A42473	ermK leader peptid
39	18	22.5	14	2	A32654	fibrinopeptide A -
40	18	22.5	15	2	S42741	ubiquinol-cytochro
41	18	22.5	15	2	S26791	Ig heavy chain V r
42	18	22.5	15	2	S10388	Ig heavy chain J r
43	18	22.5	15	2	S10386	Ig heavy chain J r
44	18	22.5	15	2	PH0780	T-cell receptor al
45	18	22.5	15	2	PC4213	bphB protein - Com

ALIGNMENTS

RESULT 1
A61512
variant surface glycoprotein MITat 1.7 - Trypanosoma brucei (fragment)
C;Species: Trypanosoma brucei
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C;Accession: A61512
R;Holder, A.A.; Cross, G.A.M.
Mol. Biochem. Parasitol. 2, 135-150, 1981
A;Title: Glycopeptides from variant surface glycoproteins of Trypanosoma brucei. C-termir
A;Reference number: A61512; MUID:81172836; PMID:6163983
A;Accession: A61512
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <HOL>
A;Cross-references: UNIPROT:Q7M3S1
C;Keywords: glycoprotein

Query Match 35.0%; Score 28; DB 2; Length 11;
Best Local Similarity 50.0%; Pred. No. 77;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy          3 GTMRTRHSTG 12
              ||  |: :||
Db          1 GTAETQBTTG 10
```

RESULT 2
PH0776
T-cell receptor alpha chain (M1 V-alpha-8.F3.3) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0776
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-re
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0776
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60873
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 28.7%; Score 23; DB 2; Length 14;
Best Local Similarity 80.0%; Pred. No. 7.6e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy         11 TGGTN 15
              ||| |
Db          5 TGGNN 9
```

```
RESULT 3
YFPG
thymic factor - pig
N;Alternate names: FTS (facteur thymique serique)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 09-Jul-2004
C;Accession: A01523; A60983
R;Pleau, J.M.; Dardenne, M.; Blouquit, Y.; Bach, J.F.
J. Biol. Chem. 252, 8045-8047, 1977
A;Title: Structural study of circulating thymic factor: a peptide isolated from pig seru
A;Reference number: A01523; MUID:78026571; PMID:914862
A;Accession: A01523
A;Molecule type: protein
A;Residues: 1-9 <PLE>
A;Cross-references: UNIPROT:P01255
R;Bach, J.F.; Dardenne, M.; Pleau, J.M.; Rosa, J.
Nature 266, 55-57, 1977
A;Title: Biochemical characterisation of a serum thymic factor.
A;Reference number: A60983; MUID:77123829; PMID:300146
A;Accession: A60983
A;Molecule type: protein
A;Residues: 'Z',2-4,'Z',6-9 <BAC>
C;Comment: This peptide induces DNA synthesis in immature thymocytes, but not peripheral
in a variety of immunoassays.
C;Comment: See PIR:A60957 (sheep) for discussion of another possible N-terminal modifica
C;Superfamily: thymic factor
C;Keywords: pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

  Query Match             27.5%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 STGGTN 15
              | ||:|
Db          4 SQGGSN 9


RESULT 4
A60957
thymocyte growth peptide - sheep
N;Contains: FTS (facteur thymique serique)
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004
C;Accession: A60957
R;Ernstroem, U.; Gafvelin, G.; Rudja, J.M.
Biosci. Rep. 10, 403-412, 1990
A;Title: Purification of thymocyte growth peptide (TGP) from sheep thymus. Relationship
A;Reference number: A60957; MUID:91064427; PMID:2249004
A;Accession: A60957
A;Molecule type: protein
A;Residues: 1-9 <ERN>
A;Cross-references: UNIPROT:Q7M3C5
C;Comment: This peptide induces DNA synthesis in immature thymocytes, but not peripheral
in a variety of immunoassays.
C;Comment: This peptide was isolated in two forms. One form contained the pyrrolidone ca
r form (thymocyte growth peptide) contains a large, non-peptide blocking group with a hi
C;Superfamily: thymic factor
C;Keywords: blocked amino end; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Glx) (in FTS) #status experimental
F;1/Modified site: blocked amino end (Glx) (in thymocyte growth peptide) #status experim

  Query Match             27.5%;  Score 22;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 STGGTN 15
              | ||:|
Db          4 SQGGSN 9


RESULT 5
PH0751
T-cell receptor beta chain (F12) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C;Accession: PH0751
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-re
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0751
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60843; NID:g50931; PIDN:CAA43235.1; PID:g50932
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             27.5%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 TGGT 14
              ||||
Db          6 TGGT 9


RESULT 6
PT0639
T-cell receptor beta chain V-D-J region (111-1AA) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0639
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0639
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-8 <FEE>
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             26.2%;  Score 21;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 STGG 13
              ||||
Db          3 STGG 6


RESULT 7
PH0919
T-cell receptor beta chain V-D-J region (isolate 5) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Oct-1992 #sequence_revision 09-Oct-1992 #text_change 30-May-1997
C;Accession: PH0919
R;Gold, D.P.; Offner, H.; Sun, D.; Wiley, S.; Vandenbark, A.A.; Wilson, D.B.
J. Exp. Med. 174, 1467-1476, 1991
A;Title: Analysis of T cell receptor beta chains in Lewis rats with experimental allergic
A;Reference number: PH0891; MUID:92078857; PMID:1836012
A;Accession: PH0919
A;Molecule type: mRNA
A;Residues: 1-11 <GOL>
A;Experimental source: concanavalin A-activated lymphoblast
A;Note: the authors translated the codon CAG for residue 11 as Glu
C;Keywords: T-cell receptor

  Query Match             26.2%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          6 RTRHSTG 12
              | | :||
Db          4 RDRRNTG 10


RESULT 8
S05002
corazonin - American cockroach
C;Species: Periplaneta americana (American cockroach)
C;Date: 07-Sep-1990 #sequence_revision 09-Apr-1998 #text_change 09-Jul-2004
C;Accession: S05002
R;Veenstra, J.A.
FEBS Lett. 250, 231-234, 1989
A;Title: Isolation and structure of corazonin, a cardioactive peptide from the american
A;Reference number: S05002; MUID:89325572; PMID:2753132
A;Accession: S05002
A;Molecule type: protein
A;Residues: 1-11 <VEE>
A;Cross-references: UNIPROT:P11496
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;11/Modified site: amidated carboxyl end (Asn) #status experimental

  Query Match             26.2%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 RHSTGGTN 15
              ::| | ||
Db          4 QYSRGWTN 11


RESULT 9
S57577
T cell receptor V-J junctional alpha chain region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57577
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A;Reference number: S57494
A;Accession: S57577
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <BUR>
A;Cross-references: EMBL:Z49945; NID:g887492; PIDN:CAA90216.1; PID:g887493
C;Keywords: T-cell receptor

  Query Match             26.2%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          8 RHSTGGTN 15
              |  ||  |
Db          4 RDQTGANN 11


RESULT 10
A54823
olfactory receptor I7 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 09-Jul-2004
C;Accession: A54823
R;Chess, A.; Simon, I.; Cedar, H.; Axel, R.
Cell 78, 823-834, 1994
A;Title: Allelic inactivation regulates olfactory receptor gene expression.
A;Reference number: A54823; MUID:94373818; PMID:8087849
A;Accession: A54823
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-8 <CHE>
A;Cross-references: UNIPROT:Q9QWU6

  Query Match             25.0%;  Score 20;  DB 2;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 MRTRHSTG 12
              |  |: ||
Db          1 MERRNHTG 8


RESULT 11
D20907
Ig kappa-1 chain J4 segment (b95 allotype) - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 10-Aug-1990 #sequence_revision 10-Aug-1990 #text_change 05-Nov-1999
C;Accession: D20907; D53275
R;Emorine, L.; Max, E.E.
Nucleic Acids Res. 11, 8877-8890, 1983
A;Title: Structural analysis of a rabbit immunoglobulin kappa2 J-C locus reveals multiple
A;Reference number: A20907; MUID:84169523; PMID:6324107
A;Accession: D20907
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-12 <EMO>
A;Cross-references: GB:X00231; NID:g1577; PIDN:CAA25049.1; PID:e8275; PID:g1364234
R;Ayadi, H.; Marche, P.N.; Cazenave, P.A.
Immunogenetics 34, 201-207, 1991
A;Title: Evolution of the rabbit immunoglobulin kappa chain genes.
A;Reference number: A53275; MUID:91372868; PMID:1909995
A;Accession: D53275
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-12 <AYA>
A;Note: sequence extracted from NCBI backbone (NCBIN:56069, NCBIP:56166)
C;Comment: This J4 segment may not be functional because of a short space between the 7 n
C;Keywords: heterotetramer; immunoglobulin

  Query Match             25.0%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 SGTM 5
              ||||
Db          5 SGTM 8


RESULT 12
PH1464
T-cell receptor beta chain (clone A3/63) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Mar-2004
C;Accession: PH1464
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Kou
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompatibi
A;Reference number: PH1430; MUID:93171821; PMID:8436911
A;Accession: PH1464
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Keywords: receptor; T-cell

  Query Match             25.0%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   80.0%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 STGGT 14
              ||| |
Db          4 STGNT 8


RESULT 13
I51432
```

histone H4-1 precursor - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51432
R;Gargiulo, G.; Razvi, F.; Ruberti, I.; Mohr, I.; Worcel, A.
J. Mol. Biol. 181, 333-349, 1985
A;Title: Chromatin-specific hypersensitive sites are assembled on a Xenopus histone gene
A;Reference number: I51431; MUID:85160855; PMID:4039007
A;Accession: I51432
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-14 <GAR>
A;Cross-references: UNIPROT:P70007; GB:M23777; NID:g214219; PIDN:AAA49737.1; PID:g214222
C;Superfamily: histone H4

  Query Match 25.0%; Score 20; DB 2; Length 14;
  Best Local Similarity 50.0%; Pred. No. 2.6e+03;
  Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 6 RTRHSTGG 13
 || : ||
Db 7 RTLYGFGG 14

RESULT 14
C43334
orf3 3' to aadR - Rhodopseudomonas palustris (fragment)
C;Species: Rhodopseudomonas palustris
C;Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: C43334
R;Dispensa, M.; Thomas, C.T.; Kim, M.K.; Perrotta, J.A.; Gibson, J.; Harwood, C.S.
J. Bacteriol. 174, 5803-5813, 1992
A;Title: Anaerobic growth of Rhodopseudomonas palustris on 4-hydroxybenzoate is dependen
A;Reference number: A43334; MUID:92394882; PMID:1522059
A;Accession: C43334
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-15 <DIS>
A;Cross-references: UNIPROT:Q02006; GB:M92426; NID:g151870; PIDN:AAA26091.1; PID:g551951
A;Note: sequence extracted from NCBI backbone (NCBIN:112964, NCBIP:112967)

  Query Match 25.0%; Score 20; DB 2; Length 15;
  Best Local Similarity 60.0%; Pred. No. 2.7e+03;
  Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 MRTRH 9
 :| ||
Db 9 LRARH 13

RESULT 15
A28497
neurotensin-related protein - turkey (fragment)
C;Species: Meleagris gallopavo (common turkey)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: A28497
R;Carraway, R.E.; Cochrane, D.E.; Ruane, S.E.
J. Biol. Chem. 262, 15886-15889, 1987
A;Title: Isolation, structures, and biologic activity of neurotensin-related peptides ge
A;Reference number: A28497; MUID:88058942; PMID:2445741
A;Accession: A28497
A;Molecule type: protein
A;Residues: 1-15 <CAR>
A;Cross-references: UNIPROT:Q7LZA3
C;Keywords: neuropeptide

  Query Match 25.0%; Score 20; DB 2; Length 15;
  Best Local Similarity 100.0%; Pred. No. 2.7e+03;
  Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 MRTR 8
 ||||
Db 7 MRTR 10

Search completed: February 22, 2005, 09:46:26
Job time : 11.1333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-7
Perfect score: 80
Sequence: 1 QSGTMRTRHSTGGTN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	35.0	11	2	Q7M3S1	Q7m3s1 trypanosoma
2	26	32.5	13	2	Q9REI2	Q9rei2 acidiphiliu
3	23	28.7	8	2	Q15894	Q15894 homo sapien
4	23	28.7	13	2	Q8I8F3	Q8i8f3 drosophila
5	23	28.7	13	2	P97944	P97944 mus musculu
6	23	28.7	13	2	Q65331	Q65331 autographa
7	23	28.7	15	2	Q95770	Q95770 cyclura ric
8	22	27.5	9	1	THYF_PIG	P01255 sus scrofa
9	22	27.5	9	2	Q7M3C5	Q7m3c5 ovis aries
10	22	27.5	12	2	Q6LDQ6	Q6ldq6 mus musculu
11	22	27.5	13	2	Q6PJ24	Q6pj24 homo sapien
12	22	27.5	14	2	Q642Z4	Q642z4 brachytheci
13	22	27.5	14	2	Q642Z5	Q642z5 brachytheci
14	22	27.5	14	2	Q642Z8	Q642z8 brachytheci
15	22	27.5	14	2	Q643A0	Q643a0 brachytheci
16	22	27.5	14	2	P83424	P83424 morus nigra
17	22	27.5	15	2	Q95762	Q95762 ctenosaura
18	22	27.5	15	2	Q6XBM6	Q6xbm6 platylomell
19	22	27.5	15	2	Q42223	Q42223 arabidopsis
20	21	26.2	8	2	Q6J0R5	Q6j0r5 paracentrot
21	21	26.2	8	2	O19956	O19956 gossypium a
22	21	26.2	8	2	O19958	O19958 gossypium b
23	21	26.2	8	2	O19959	O19959 gossypium t
24	21	26.2	8	2	O19960	O19960 gossypium m
25	21	26.2	8	2	O19961	O19961 gossypium d
26	21	26.2	10	2	Q8SAC2	Q8sac2 amblystegiu
27	21	26.2	11	1	CORZ_PERAM	P11496 periplaneta
28	21	26.2	14	2	Q79AV9	Q79av9 klebsiella
29	20	25.0	8	2	Q80WD6	Q80wd6 mus musculu
30	20	25.0	12	2	Q841R5	Q841r5 agrobacteri
31	20	25.0	13	2	Q7RDM8	Q7rdm8 plasmodium
32	20	25.0	13	2	P97140	P97140 borrelia bu
33	20	25.0	13	2	Q53693	Q53693 streptomyce
34	20	25.0	14	2	P70007	P70007 xenopus lae
35	20	25.0	15	2	Q9TRP3	Q9trp3 sus scrofa
36	20	25.0	15	2	Q7LZA3	Q7lza3 meleagris g
37	19	23.8	8	2	Q84156	Q84156 orf virus.
38	19	23.8	8	2	Q84271	Q84271 human papil
39	19	23.8	9	1	FARD_CALVO	P41868 calliphora
40	19	23.8	9	2	Q6LEH2	Q6leh2 homo sapien
41	19	23.8	10	2	Q7M3S4	Q7m3s4 trypanosoma
42	19	23.8	10	2	Q7M278	Q7m278 triticum tu
43	19	23.8	11	2	Q9BJ61	Q9bj61 dictyosteli
44	19	23.8	12	2	Q6DUJ5	Q6duj5 rattus norv
45	19	23.8	15	2	Q9UWG1	Q9uwg1 pyrococcus,

ALIGNMENTS

RESULT 1
Q7M3S1
ID Q7M3S1 PRELIMINARY; PRT; 11 AA.
AC Q7M3S1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Variant surface glycoprotein MITat 1.7 (Fragment).
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE.
RX MEDLINE=81172836; PubMed=6163983; DOI=10.1016/0166-6851(81)90095-5;
RA Holder A.A., Cross G.A.M.;
RT "Glycopeptides from variant surface glycoproteins of Trypanosoma
RT brucei. C-terminal location of antigenically cross-reacting
RT carbohydrate moieties.";
RL Mol. Biochem. Parasitol. 2:135-150(1981).
DR PIR; A61512; A61512.
FT NON_TER 1 1
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1067 MW; 64FD322AAF16C1AB CRC64;

Query Match 35.0%; Score 28; DB 2; Length 11;
Best Local Similarity 50.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy          3 GTMRTRHSTG 12
              ||  |: :||
Db          1 GTAETQBTTG 10
```

RESULT 2
Q9REI2
ID Q9REI2 PRELIMINARY; PRT; 13 AA.
AC Q9REI2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Acidiphilium symbioticum.
OG Plasmid pAS3.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Acidiphilium.
OX NCBI_TaxID=94005;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KM2;
RX MEDLINE=22165970; PubMed=12177739;
RA Mahapatra N.R., Ghosh S., Deb C., Banerjee P.C.;
RT "Resistance to cadmium and zinc in Acidiphilium symbioticum KM2 is
RT plasmid mediated.";
RL Curr. Microbiol. 45:180-186(2002).

```
DR   EMBL; AJ239066; CAB65725.1; -.
KW   Hypothetical protein; Plasmid.
SQ   SEQUENCE   13 AA;  1361 MW;  6811778519FD11B4 CRC64;

  Query Match             32.5%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   55.6%;  Pred. No. 8.5e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 MRTRHSTGG 13
              :||| : ||
Db          5 IRTRAARGG 13


RESULT 3
Q15894
ID   Q15894      PRELIMINARY;      PRT;     8 AA.
AC   Q15894;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Homo sapiens (clone XP587B) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RA   Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA   Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA   Caskey C.T.H.;
RT   "Isolation of chromosome-specific genes by reciprocal probing of
RT   arrayed cDNAs and cosmid libraries.";
RL   Hum. Mol. Genet. 0:0-0(1995).
DR   EMBL; L32074; AAA73884.1; -.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  952 MW;  EBC735B1E1F1B6D6 CRC64;

  Query Match             28.7%;  Score 23;  DB 2;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 MRTRHS 10
              |:| ||
Db          1 MQTHHS 6


RESULT 4
Q8I8F3
ID   Q8I8F3      PRELIMINARY;      PRT;    13 AA.
AC   Q8I8F3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Acetylcholinesterase (Fragment).
GN   Name=Ace; Synonyms=ache;
OS   Drosophila parisiena.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila; mayaguana subcluster.
OX   NCBI_TaxID=214822;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   O'Grady P.M. II, Durando C.M., Heed W.B., Wasserman M., Etges W.,
RA   DeSalle R.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY154986; AAO13634.1; -.
DR   FlyBase; FBgn0064412; Dpsi\Ace.
FT   NON_TER       1       1
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1451 MW;  67DAF26C3079B774 CRC64;

  Query Match             28.7%;  Score 23;  DB 2;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 SGTMRTRHST 11
              ||| |   :|
Db          2 SGTSRPTRTT 11


RESULT 5
P97944
ID   P97944      PRELIMINARY;      PRT;    13 AA.
AC   P97944;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Calcium channel beta subunit (Fragment).
GN   Name=Cacnb4; Synonyms=Cchb4;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=lethargic, and BALB/c;
RX   MEDLINE=97191307; PubMed=9039265; DOI=10.1016/S0092-8674(00)81877-2;
RA   Burgess D.L., Jones J.M., Meisler M.H., Noebels J.L.;
RT   "Mutation of the Ca2+ channel beta subunit gene Cchb4 is associated
RT   with ataxia and seizures in the lethargic (lh) mouse.";
RL   Cell 88:385-392(1997).
DR   EMBL; U80986; AAC53037.1; -.
DR   EMBL; U80985; AAC53036.1; -.
DR   MGD; MGI:103301; Cacnb4.
DR   GO; GO:0005891; C:voltage-gated calcium channel complex; TAS.
DR   GO; GO:0008331; F:high voltage-gated calcium channel activity; TAS.
FT   NON_TER       1       1
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1370 MW;  DFE26C2AA094C409 CRC64;

  Query Match             28.7%;  Score 23;  DB 2;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 SGTMRTRHST 11
              ||| |   :|
Db          3 SGTFRATPTT 12


RESULT 6
Q65331
ID   Q65331      PRELIMINARY;      PRT;    13 AA.
AC   Q65331;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Basic protein (Fragment).
OS   Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC   Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC   Nucleopolyhedrovirus.
OX   NCBI_TaxID=46015;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6;
RX   MEDLINE=90218045; PubMed=2109042;
RA   Hill-Perkins M.S., Possee R.D.;
RT   "A baculovirus expression vector derived from the basic protein
RT   promoter of Autographa californica nuclear polyhedrosis virus.";
RL   J. Gen. Virol. 71:971-976(1990).
DR   EMBL; D00864; BAA00738.1; -.
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1626 MW;  6C75B58A749D1414 CRC64;

  Query Match             28.7%;  Score 23;  DB 2;  Length 13;
```

```
  Best Local Similarity   71.4%;  Pred. No. 2.9e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 RTRHSTG 12
              | | |||
Db          6 RRRSSTG 12


RESULT 7
Q95770
ID   Q95770      PRELIMINARY;      PRT;    15 AA.
AC   Q95770;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   NADH dehydrogenase subunit 4 (Fragment).
GN   Name=ND4;
OS   Cyclura ricordi (Ricord's rock iguana).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Iguania; Iguanidae; Iguaninae; Cyclura.
OX   NCBI_TaxID=51215;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97019047; PubMed=8865663;
RA   Sites J.W. Jr., Davis S.K., Guerra T., Iverson J.B., Snell H.L.;
RT   "Character congruence and phylogenetic signal in molecular and
RT   morphological data sets: a case study in the living Iguanas (Squamata,
RT   Iguanidae).";
RL   Mol. Biol. Evol. 13:1087-1105(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sites J.W., Davis S.K., Guerra T., Iverson J.B., Snell H.L.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U66237; AAB07473.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1715 MW;  8327178E7927A57E CRC64;

  Query Match             28.7%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 TRHST 11
              :||||
Db          9 SRHST 13


RESULT 8
THYF_PIG
ID   THYF_PIG       STANDARD;      PRT;     9 AA.
AC   P01255;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Thymic factor.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=78026571; PubMed=914862;
RA   Pleau J.-M., Dardenne M., Blouquit Y., Bach J.-F.;
RT   "Structural study of circulating thymic factor: a peptide isolated
RT   from pig serum. II. Amino acid sequence.";
RL   J. Biol. Chem. 252:8045-8047(1977).
CC   -!- MISCELLANEOUS: The biological source(s) and physiological
CC       activities of this peptide have not been determined.
DR   PIR; A01523; YFPG.
KW   Direct protein sequencing; Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       Pyrrolidone carboxylic acid.
SQ   SEQUENCE   9 AA;  876 MW;  D500B87866C5B33D CRC64;

  Query Match             27.5%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 STGGTN 15
              | ||:|
Db          4 SQGGSN 9


RESULT 9
Q7M3C5
ID   Q7M3C5      PRELIMINARY;      PRT;     9 AA.
AC   Q7M3C5;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Thymocyte growth peptide.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE.
RA   Ernstroem U., Gafvelin G., Rudja J.M.;
RT   "Purification of thymocyte growth peptide (TGP) from sheep thymus.
RT   Relationship to FTS/thymulin.";
RL   Biosci. Rep. 10:403-412(1990).
DR   PIR; A60957; A60957.
SQ   SEQUENCE   9 AA;  876 MW;  DCE7B87866C5B33D CRC64;

  Query Match             27.5%;  Score 22;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 STGGTN 15
              | ||:|
Db          4 SQGGSN 9


RESULT 10
Q6LDQ6
ID   Q6LDQ6      PRELIMINARY;      PRT;    12 AA.
AC   Q6LDQ6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   C-myb oncogene, exon 1 and exon 2. (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90017472; PubMed=2678098;
RA   Reddy C.D., Reddy P.E.;
RT   "Differential binding of nuclear factors to intron-1 containing the
RT   transcriptional pause site correlates with c-myb expression.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:7326-7330(1989).
DR   EMBL; M26185; AAA37505.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1461 MW;  32769A1F9CB1F417 CRC64;

  Query Match             27.5%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 4e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 RTRHST 11
              | |||:
Db          4 RPRHSS 9
```

```
RESULT 11
Q6PJ24
ID   Q6PJ24      PRELIMINARY;      PRT;    13 AA.
AC   Q6PJ24;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   POLRMT protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Strausberg R.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC024170; AAH24170.1; -.
SQ   SEQUENCE   13 AA;  1266 MW;  7D6737769F9D0734 CRC64;

  Query Match             27.5%;  Score 22;  DB 2;  Length 13;
  Best Local Similarity   36.4%;  Pred. No. 4.4e+03;
  Matches    4;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 SGTMRTRHSTG 12
              :|  | |  ::|
Db          2 AGACRLRSASG 12


RESULT 12
Q642Z4
ID   Q642Z4      PRELIMINARY;      PRT;    14 AA.
AC   Q642Z4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   AtpB (Fragment).
GN   Name=atpB;
OS   Brachytheciastrum venustum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC   Bryopsida; Bryidae; Hypnales; Brachytheciaceae; Brachytheciastrum.
OX   NCBI_TaxID=292614;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Venustum1704, and Venustum1705;
RA   Vanderpoorten A.;
RT   "A molecular and morphological recircumscription of Brachytheciastrum
RT   (Brachytheciaceae, Bryopsida).";
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY736270; AAU25912.1; -.
DR   EMBL; AY736271; AAU25913.1; -.
KW   Chloroplast.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1544 MW;  811BBAA1FD5040D1 CRC64;

  Query Match             27.5%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   45.5%;  Pred. No. 4.8e+03;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 MRTRHSTGGTN 15
              |:|   | ||:
Db          1 MKTDSRTFGTS 11


RESULT 13
Q642Z5
ID   Q642Z5      PRELIMINARY;      PRT;    14 AA.
AC   Q642Z5;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   AtpB (Fragment).
GN   Name=atpB;
OS   Brachytheciastrum velutinum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC   Bryopsida; Bryidae; Hypnales; Brachytheciaceae; Brachytheciastrum.
OX   NCBI_TaxID=113273;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Velu1771;
RA   Vanderpoorten A.;
RT   "A molecular and morphological recircumscription of Brachytheciastrum
RT   (Brachytheciaceae, Bryopsida).";
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY736269; AAU25911.1; -.
KW   Chloroplast.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1544 MW;  811BBAA1FD5040D1 CRC64;

  Query Match             27.5%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   45.5%;  Pred. No. 4.8e+03;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 MRTRHSTGGTN 15
              |:|   | ||:
Db          1 MKTDSRTFGTS 11


RESULT 14
Q642Z8
ID   Q642Z8      PRELIMINARY;      PRT;    14 AA.
AC   Q642Z8;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   AtpB (Fragment).
GN   Name=atpB;
OS   Brachytheciastrum trachypodium.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC   Bryopsida; Bryidae; Hypnales; Brachytheciaceae; Brachytheciastrum.
OX   NCBI_TaxID=292612;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Trachypo1762, Trachypo1766, and Trachypo1770;
RA   Vanderpoorten A.;
RT   "A molecular and morphological recircumscription of Brachytheciastrum
RT   (Brachytheciaceae, Bryopsida).";
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
```

```
DR   EMBL; AY736266; AAU25908.1; -.
DR   EMBL; AY736267; AAU25909.1; -.
DR   EMBL; AY736268; AAU25910.1; -.
KW   Chloroplast.
FT   NON_TER       14      14
SQ   SEQUENCE   14 AA;  1544 MW;  811BBAA1FD5040D1 CRC64;

  Query Match             27.5%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   45.5%;  Pred. No. 4.8e+03;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 MRTRHSTGGTN 15
              |:|   | ||:
Db          1 MKTDSRTFGTS 11


RESULT 15
Q643A0
ID   Q643A0      PRELIMINARY;      PRT;    14 AA.
AC   Q643A0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   AtpB (Fragment).
GN   Name=atpB;
OS   Brachytheciastrum collinum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC   Bryopsida; Bryidae; Hypnales; Brachytheciaceae; Brachytheciastrum.
OX   NCBI_TaxID=219719;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Collinum1767, Collinum1769, and Collinum1765;
RA   Vanderpoorten A.;
RT   "A molecular and morphological recircumscription of Brachytheciastrum
RT   (Brachytheciaceae, Bryopsida).";
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY736264; AAU25906.1; -.
DR   EMBL; AY736265; AAU25907.1; -.
DR   EMBL; AY736263; AAU25905.1; -.
KW   Chloroplast.
FT   NON_TER       14      14
SQ   SEQUENCE   14 AA;  1544 MW;  811BBAA1FD5040D1 CRC64;

  Query Match             27.5%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   45.5%;  Pred. No. 4.8e+03;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 MRTRHSTGGTN 15
              |:|   | ||:
Db          1 MKTDSRTFGTS 11


Search completed: February 22, 2005, 09:37:56
Job time : 54.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-7
Perfect score: 80
Sequence: 1 QSGTMRTRHSTGGTN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	80	100.0	15	2	AAW04847	Aaw04847 Self epit
2	41	51.2	14	4	ABB56600	Abb56600 Human SNP
3	34	42.5	10	4	AAG86787	Aag86787 Saccharom
4	31	38.8	12	2	AAR13070	Aar13070 SIV500 fu
5	30	37.5	15	7	ADM37098	Adm37098 Mutated I
6	30	37.5	15	7	ADM53386	Adm53386 Mutant NF
7	30	37.5	15	7	ADM53392	Adm53392 Mutant NF
8	30	37.5	15	7	ADM95020	Adm95020 Converted
9	30	37.5	15	8	ADR31520	Adr31520 IkappaB-a
10	30	37.5	15	8	ADR31534	Adr31534 IkappaB-a
11	29	36.2	7	2	AAW59304	Aaw59304 Non-polio
12	29	36.2	7	2	AAY50064	Aay50064 Coxsackie
13	29	36.2	11	3	AAB26509	Aab26509 Human IgE
14	29	36.2	11	4	AAU16840	Aau16840 Peptide P
15	29	36.2	11	5	ABJ00284	Abj00284 Human IgE
16	29	36.2	14	3	AAB26512	Aab26512 Human IgE
17	29	36.2	14	4	AAM00598	Aam00598 Human pro
18	29	36.2	14	4	AAU16843	Aau16843 Peptide P
19	29	36.2	14	5	ABJ00287	Abj00287 Human IgE
20	28	35.0	7	5	ABP49196	Abp49196 Zinc fing
21	28	35.0	7	7	ADA63567	Ada63567 Zinc fing
22	28	35.0	7	7	ADM22065	Adm22065 Synthetic
23	28	35.0	8	7	ADF66168	Adf66168 Mouse ubi
24	28	35.0	10	4	AAG63152	Aag63152 Peptide d
25	28	35.0	10	4	AAG97581	Aag97581 Human com
26	28	35.0	10	4	AAG97582	Aag97582 Human com
27	28	35.0	13	3	AAY66866	Aay66866 T cell an
28	28	35.0	13	5	ADG67012	Adg67012 Ricin tox
29	28	35.0	13	5	ADG67013	Adg67013 Ricin tox
30	28	35.0	13	5	ADG67014	Adg67014 Ricin tox
31	28	35.0	13	5	ADG67015	Adg67015 Ricin tox
32	28	35.0	14	3	ADC16855	Adc16855 Human sin
33	28	35.0	14	6	ABR75388	Abr75388 Biologica
34	27.5	34.4	13	2	AAR77321	Aar77321 Human apo
35	27	33.8	7	3	AAB02885	Aab02885 Nucleotid
36	27	33.8	7	5	ABP49458	Abp49458 Zinc fing
37	27	33.8	7	5	ABP49464	Abp49464 Zinc fing
38	27	33.8	7	5	ABP50954	Abp50954 Zinc fing
39	27	33.8	7	5	ABP48644	Abp48644 Zinc fing
40	27	33.8	7	5	ABP49467	Abp49467 Zinc fing
41	27	33.8	7	5	ABP48647	Abp48647 Zinc fing
42	27	33.8	7	5	ABP49461	Abp49461 Zinc fing
43	27	33.8	7	7	ADA63126	Ada63126 Zinc fing
44	27	33.8	7	7	ADA63760	Ada63760 Zinc fing
45	27	33.8	7	7	ADA63758	Ada63758 Zinc fing

ALIGNMENTS

RESULT 1
AAW04847
ID AAW04847 standard; peptide; 15 AA.
XX
AC AAW04847;
XX
DT 18-FEB-1997 (first entry)
XX
DE Self epitope of desmoglein 3, implicated in autoimmune disease.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Homo sapiens.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 1; Page 42; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is derived from the human desmoglein

```
CC   3 protein (amino acids 762-786) and is implicated as a self epitope in
CC   pemphigus vulgaris. Peptides derived from the human desmoglein protein
CC   are described in AAW04841-47
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 80;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 6e-07;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QSGTMRTRHSTGGTN 15
              |||||||||||||||
Db          1 QSGTMRTRHSTGGTN 15


RESULT 2
ABB56600
ID   ABB56600 standard; peptide; 14 AA.
XX
AC   ABB56600;
XX
DT   05-MAR-2002  (first entry)
XX
DE   Human SNP related amino acid sequence SEQ ID NO:1165.
XX
KW   Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW   immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW   autoimmune disease; inflammation; cancer; nervous system disease;
KW   infection; polymorphic protein.
XX
OS   Homo sapiens.
XX
PN   WO200138586-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032311.
XX
PR   24-NOV-1999;   99US-0167383P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-355949/37.
XX
PT   Isolated human nucleic acids comprising one or more single nucleotide
PT   polymorphisms, useful for treating a subject suffering from a pathology,
PT   e.g. autoimmune diseases, ascribed to the presence of a sequence
PT   polymorphism.
XX
PS   Claim 1; Page 600; 674pp; English.
XX
CC   ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC   comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC   to ABB56903 represent human peptides encoded by some of the SNP
CC   oligonucleotides. The sequences from the present invention can have
CC   immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC   antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC   and antibodies from the present invention can be used for treating a
CC   subject suffering from, at risk for, or suspected of, suffering from a
CC   pathology ascribed to the presence of a sequence polymorphism. The
CC   pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC   the nervous system, and infection by pathogenic microorganisms. The SNPs
CC   are also useful for determining which forms of a characterised
CC   polymorphism are present in individuals. The antibodies may be used in
CC   the detection, quantitation and/or cellular or tissue localisation of a
CC   polymorphic protein (e.g., for use in measuring levels of the polymorphic
CC   protein within appropriate physiological samples)
XX
SQ   Sequence 14 AA;

  Query Match             51.2%;  Score 41;  DB 4;  Length 14;
  Best Local Similarity   70.0%;  Pred. No. 3.4;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 RTRHSTGGTN 15
              : || |||||
Db          1 QARHRTGGTN 10


RESULT 3
AAG86787
ID   AAG86787 standard; peptide; 10 AA.
XX
AC   AAG86787;
XX
DT   11-SEP-2001  (first entry)
XX
DE   Saccharomyces cerevisiae peptide, SEQ ID NO: 1736.
XX
KW   Saccharomyces cerevisiae; complementary peptide; peptide identification;
KW   drug discovery; drug design.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO200142276-A1.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB004773.
XX
PR   13-DEC-1999;   99GB-00029471.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-367863/38.
XX
PT   Identifying complementary peptides by analysis of protein and nucleotide
PT   sequence databases, useful in drug design.
XX
PS   Example 3; Page 269; 488pp; English.
XX
CC   The invention relates to the identification of complementary peptides by
CC   analysis of protein and nucleotide sequence databases from higher
CC   eukaryotic genomes, excluding human and plants. The specific
CC   complementary peptides interact with their relevant target proteins
CC   encoded in the eukaryote genome. The peptides may be used as reagents and
CC   drugs for drug discovery and as lead ligands for drug design and
CC   development. The present sequence is a complementary peptide from
CC   Saccharomyces cerevisiae
XX
SQ   Sequence 10 AA;

  Query Match             42.5%;  Score 34;  DB 4;  Length 10;
  Best Local Similarity   70.0%;  Pred. No. 38;
  Matches    7;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 RTRHSTGGTN 15
              ||| | || |
Db          1 RTRASRGGAN 10


RESULT 4
AAR13070
ID   AAR13070 standard; protein; 12 AA.
XX
AC   AAR13070;
XX
DT   04-OCT-1991  (first entry)
XX
DE   SIV500 fusion protein.
XX
KW   SIV; beta-galactosidase; vaccine; immunotoxin; antibodies.
```

XX
OS Synthetic.
XX
PN WO9109872-A.
XX
PD 11-JUL-1991.
XX
PF 13-DEC-1989; 89US-00450150.
XX
PR 13-DEC-1989; 89US-00450150.
XX
PA (UNIV-) UNIVAX BIOLOGICS IN.
XX
PI Shafferman A;
XX
DR WPI; 1991-222846/30.
DR N-PSDB; AAQ12645.
XX
PT New fusion proteins for diagnosis and treatment of HIV infection -
PT comprise synthetic antigenic determinants of HIV fused to non-HIV
PT polypeptide, preferably beta galactosidase.
XX
PS Disclosure; Fig 2B; 58pp; English.
XX
CC The sequence identified from SIVmac env is equivalent to HIV500
CC (AAQ12641). Because of the criteria used to correlate HIV and SIV
CC sequences, observation of the antibody response of monkeys immunised with
CC an SIV sequence could be used to assess the antibody response of humans
CC immunised with the corresp. HIV sequence. See also AAQ12640-47
XX
SQ Sequence 12 AA;

Query Match 38.8%; Score 31; DB 2; Length 12;
Best Local Similarity 62.5%; Pred. No. 1.6e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 8 RHSTGGTN 15
|::||||:
Db 1 RYTTGGTS 8

RESULT 5
ADM37098
ID ADM37098 standard; peptide; 15 AA.
XX
AC ADM37098;
XX
DT 03-JUN-2004 (first entry)
XX
DE Mutated IkappaB-alpha 26-40.
XX
KW IkappaB-alpha; AKT kinase; beta-TrCP; kinase; phosphorylation substrate;
KW kinase recognition domain; E3 binding region; mutant; mutein.
XX
OS Synthetic.
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 2
FT /note= "Wild-type Asp substituted by Arg"
FT Misc-difference 7
FT /note= "Wild-type Ser substituted by Thr"
XX
PN US2003170611-A1.
XX
PD 11-SEP-2003.
XX
PF 09-MAR-2002; 2002US-00093840.
XX
PR 09-MAR-2002; 2002US-00093840.
XX
PA (CARD/) CARDONE M H.
PA (YAFF/) YAFFE M.
XX
PI Cardone MH, Yaffe M;
XX
DR WPI; 2003-863752/80.
XX
PT Identifying a molecule capable of modulating a kinase activity in situ
PT using a substrate with an altered kinase recognition domain and an
PT associated label is useful in drug discovery.
XX
PS Disclosure; Fig 3a; 25pp; English.
XX
CC The invention relates to identifying a molecule capable of modulating a
CC kinase activity in situ, comprising exposing a candidate molecule to a
CC cell comprising a phosphorylation substrate associated with a detectable
CC label and having a kinase recognition domain altered to be recognised by
CC a kinase that does not recognise the substrate in its unaltered state,
CC and determining whether the candidate molecule causes a change in an
CC expression of the label. Also included are a molecule capable of
CC modulating activity of at least one kinase in situ identified by the
CC claimed method, a fusion protein (comprising an E3 binding region, a
CC kinase recognition domain and a green fluorescent protein, where the
CC kinase recognition domain is the domain of beta-catenin, HIV protein VPU,
CC p27, Bcl-2 or c-Jun), a fusion protein (comprising an E3 binding region,
CC a kinase recognition domain and an enzyme capable of producing a
CC detectable enzymatic product, where the kinase recognition domain is the
CC domain of beta-catenin, HIV protein VPU, p27, Bcl-2 or c-Jun), an
CC isolated genetic molecule encoding one of the above fusion proteins, a
CC vector capable of expressing the above genetic molecule and a cell
CC transfected with the above vector. Expression of the label requires
CC phosphorylation of the phosphorylation substrate by the kinase. The
CC kinase recognition domain is altered to include a consensus recognition
CC motif for the kinase (e.g. AKT kinase). The invention is useful to study
CC kinase activity in situ and to screen for molecules that modulate kinase
CC activities in situ, for example in drug discovery. The invention allows
CC for information on multiple kinases to be provided simultaneously, which
CC prior art does not provide. The present sequence is a mutated Beta-TrCP
CC binding region of IkappaB-alpha, which is mutated to an AKT or IKK (both
CC not defined) phosphorylation site.
XX
SQ Sequence 15 AA;

Query Match 37.5%; Score 30; DB 7; Length 15;
Best Local Similarity 62.5%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 MRTRHSTG 12
:| || ||
Db 1 LRDRHDTG 8

RESULT 6
ADM53386
ID ADM53386 standard; peptide; 15 AA.
XX
AC ADM53386;
XX
DT 03-JUN-2004 (first entry)
XX
DE Mutant NFkB-regulating protein IkB-alpha IKK recogntion #1.
XX
KW kinase activity modulator; signaling enzyme; drug design; anti-tumour;
KW anti-inflammation; anti-ischaemia;
KW necrosis factor kappa B regulating protein; NFkB-regulating protein;
KW I kappa B-alpha; cytokine inducible IkappaB kinase recogntion site;
KW IKK recognition site; phosphorylation; serine/threonine kinase; AKT;
KW mutant; mutein.
XX
OS Unidentified.
OS Synthetic.
XX
PN US2003170737-A1.
XX
PD 11-SEP-2003.

XX
PF 09-MAR-2002; 2002US-00093945.
XX
PR 09-MAR-2002; 2002US-00093945.
XX
PA (CARD/) CARDONE M H.
PA (YAFF/) YAFFE M.
XX
PI Cardone MH, Yaffe M;
XX
DR WPI; 2003-787546/74.
XX
PT A method for identifying a molecule capable of modulating a kinase
PT activity in situ and is useful in drug design to screen for molecules
PT that are candidates for anti-tumor, anti-inflammation, and anti-ischemia
PT therapy.
XX
PS Disclosure; Fig 3A; 25pp; English.
XX
CC The invention describes a method of identifying a molecule capable of
CC modulating a kinase activity in situ, comprising exposing a candidate
CC molecule to a cell comprising a signaling enzyme altered so as to bind a
CC phosphorylation substrate, where the substrate is associated with
CC detectable label and the binding is regulated by a kinase, and
CC determining whether the candidate causes a change in expression of the
CC label. Also described are: a molecule capable of modulating a kinase
CC activity in situ identified by the claimed method; a fusion protein
CC comprising a genetically altered signaling enzyme and a label, where the
CC alteration produces an adapter molecule in the signaling enzyme capable
CC of binding to a phosphorylation substrate that the enzyme does not bind
CC in its unaltered state, where the binding is regulated by a kinase; an
CC isolated genetic molecule encoding the above fusion protein; a vector
CC capable of expressing the above genetic molecule; and a cell transfected
CC with the above vector. The invention is useful in drug design to screen
CC for molecules that are candidates for anti-tumour, anti-inflammation, and
CC anti-ischaemia therapy. This is the amino acid sequence of a necrosis
CC factor kappa B (NFkB)-regulating protein I kappa B-alpha cytokine
CC inducible IkappaB kinase (IKK) recognition site that has been altered to
CC allow phosphorylation by serine/threonine kinase AKT.
XX
SQ Sequence 15 AA;

Query Match 37.5%; Score 30; DB 7; Length 15;
Best Local Similarity 62.5%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 MRTRHSTG 12
:| || ||
Db 1 LRDRHDTG 8

RESULT 7
ADM53392
ID ADM53392 standard; peptide; 15 AA.
XX
AC ADM53392;
XX
DT 03-JUN-2004 (first entry)
XX
DE Mutant NFkB-regulating protein IkB-alpha IKK recogntion #2.
XX
KW kinase activity modulator; signaling enzyme; drug design; anti-tumour;
KW anti-inflammation; anti-ischaemia;
KW necrosis factor kappa B regulating protein; NFkB-regulating protein;
KW I kappa B-alpha; cytokine inducible IkappaB kinase recogntion site;
KW IKK recognition site; phosphorylation; serine/threonine kinase; AKT;
KW mutant; mutein.
XX
OS Unidentified.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 11
FT /label= Asp, Glu
XX
PN US2003170737-A1.
XX
PD 11-SEP-2003.
XX
PF 09-MAR-2002; 2002US-00093945.
XX
PR 09-MAR-2002; 2002US-00093945.
XX
PA (CARD/) CARDONE M H.
PA (YAFF/) YAFFE M.
XX
PI Cardone MH, Yaffe M;
XX
DR WPI; 2003-787546/74.
XX
PT A method for identifying a molecule capable of modulating a kinase
PT activity in situ and is useful in drug design to screen for molecules
PT that are candidates for anti-tumor, anti-inflammation, and anti-ischemia
PT therapy.
XX
PS Disclosure; Page 7; 25pp; English.
XX
CC The invention describes a method of identifying a molecule capable of
CC modulating a kinase activity in situ, comprising exposing a candidate
CC molecule to a cell comprising a signaling enzyme altered so as to bind a
CC phosphorylation substrate, where the substrate is associated with
CC detectable label and the binding is regulated by a kinase, and
CC determining whether the candidate causes a change in expression of the
CC label. Also described are: a molecule capable of modulating a kinase
CC activity in situ identified by the claimed method; a fusion protein
CC comprising a genetically altered signaling enzyme and a label, where the
CC alteration produces an adapter molecule in the signaling enzyme capable
CC of binding to a phosphorylation substrate that the enzyme does not bind
CC in its unaltered state, where the binding is regulated by a kinase; an
CC isolated genetic molecule encoding the above fusion protein; a vector
CC capable of expressing the above genetic molecule; and a cell transfected
CC with the above vector. The invention is useful in drug design to screen
CC for molecules that are candidates for anti-tumour, anti-inflammation, and
CC anti-ischaemia therapy. This is the amino acid sequence of a necrosis
CC factor kappa B (NFkB)-regulating protein I kappa B-alpha cytokine
CC inducible IkappaB kinase (IKK) recognition site that has been altered to
CC stop IkB-alpha phosphorylation.
XX
SQ Sequence 15 AA;

Query Match 37.5%; Score 30; DB 7; Length 15;
Best Local Similarity 62.5%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 MRTRHSTG 12
:| || ||
Db 1 LRDRHDTG 8

RESULT 8
ADM95020
ID ADM95020 standard; peptide; 15 AA.
XX
AC ADM95020;
XX
DT 17-JUN-2004 (first entry)
XX
DE Converted I kappa B (IkB) kinase recognition domain.
XX
KW drug design; anti-inflammatory; anti-tumour; immune response;
KW anti-ischaemic; I kappa B kinase; IkB kinase.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 11

```
FT                   /note= "Optionally Asp or Glu"
XX
PN   US2003170850-A1.
XX
PD   11-SEP-2003.
XX
PF   09-MAR-2002; 2002US-00093838.
XX
PR   09-MAR-2002; 2002US-00093838.
XX
PA   (CARD/) CARDONE M H.
PA   (YAFF/) YAFFE M.
XX
PI   Cardone MH,  Yaffe M;
XX
DR   WPI; 2003-863758/80.
XX
PT   A fusion protein comprising a pair of sequestering motifs, an enzyme-
PT   binding region and a label is useful to test and design drugs e.g. anti-
PT   inflammatory, anti-tumor, immune response regulating and anti-ischemic
PT   drugs.
XX
PS   Disclosure; Fig 3A; 24pp; English.
XX
CC   The invention relates to a fusion protein comprising a pair of
CC   sequestering motifs, an enzyme-binding region and a label, where the
CC   sequestering motifs flank each side of the enzyme-binding region which is
CC   prevented from biding to a signalling enzyme when the motifs interact
CC   with each other and the interaction is regulated by a kinase. The
CC   sequestering motifs comprise a phosphorylation substrate for the kinase
CC   and a binding motif for the phosphorylation substrate. Interaction
CC   between the phosphorylation substrate and the binding motif requires
CC   phosphorylation of the substrate by the kinase. The binding motif
CC   comprises substantially a ligand-binding region of a 14-3-3 protein or an
CC   SH2 domain and an adapter molecule which is a PDZ, SH3, WW, PTB or FHA
CC   domain. The signalling enzyme is a ubiquitin E3 ligase and the enzyme-
CC   binding region is an E3-binding region, preferably comprising a PEST
CC   element or a cyclin destruction box. Alternatively the signalling enzyme
CC   is a transporting protein. The label is fluorescent, preferably green
CC   fluorescent protein or an enzyme, preferably beta-galactosidase, a
CC   phosphatase or luciferase. The method is useful to test and design drugs
CC   with various clinical applications including anti-inflammatory candidate
CC   molecules, anti-tumour therapies, immune response regulating drugs and
CC   anti-ischaemic drugs. The present sequence represents an enzyme binding
CC   region used in the fusion protein of the invention.
XX
SQ   Sequence 15 AA;

  Query Match             37.5%;  Score 30;  DB 7;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 3e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 MRTRHSTG 12
              :| || ||
Db          1 LRDRHDTG 8


RESULT 9
ADR31520
ID   ADR31520 standard; peptide; 15 AA.
XX
AC   ADR31520;
XX
DT   04-NOV-2004  (first entry)
XX
DE   IkappaB-alpha mutant substrate peptide #1.
XX
KW   Drug designing; clinical application; anti-inflammatory; anti-tumour;
KW   immune response; anti-ischaemic; artherosclerosis; therapy;
KW   IkappaB-alpha; mutant; mutein.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 2
FT                   /note= "Wild type Asp substituted with Arg"
FT   Binding-site    6. .11
FT                   /note= "Beta-TrCP binding region 11"
FT   Misc-difference 7
FT                   /note= "Wild type Ser substituted with Thr"
FT                   /note= "Phosphorylated"
FT   Modified-site   7
FT                   /note= "Phosphorylated"
FT   Modified-site   11
FT                   /note= "Phosphorylated"
XX
PN   US2004157272-A1.
XX
PD   12-AUG-2004.
XX
PF   03-FEB-2004; 2004US-00771035.
XX
PR   09-MAR-2002; 2002US-00093945.
XX
PA   (MERR-) MERRIMACK PHARM INC.
XX
PI   Cardone MH,  Yaffe M;
XX
DR   WPI; 2004-580267/56.
XX
PT   Identifying molecule capable of modulating kinase activity in situ, by
PT   exposing candidate molecule to cell comprising signaling enzyme altered
PT   to bind substrate having label and determining whether molecule changes
PT   expression of label.
XX
PS   Disclosure; Fig 3A; 25pp; English.
XX
CC   The invention relates to a method for identifying molecule capable of
CC   modulating kinase activity in situ. The method involves exposing
CC   candidate molecule to cell having signalling enzyme altered to bind
CC   phosphorylation substrate associated with detectable label where the
CC   binding between altered signalling enzyme and substrate is regulated by a
CC   kinase and determining if the candidate molecule cause a change in
CC   expression of label. The method is useful for testing and designing drugs
CC   with various clinical application e.g. anti-inflammatory candidate
CC   molecules. The invention is useful in the field of anti-tumour
CC   therapeutics and immune response regulating drugs. The method is also
CC   useful for developing anti-ischaemic drugs that are useful for treating
CC   artherosclerosis. The present sequence is IkappaB-alpha mutant substrate
CC   peptide. This sequence is used to illustrate the method of the invention.
XX
SQ   Sequence 15 AA;

  Query Match             37.5%;  Score 30;  DB 8;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 3e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 MRTRHSTG 12
              :| || ||
Db          1 LRDRHDTG 8


RESULT 10
ADR31534
ID   ADR31534 standard; peptide; 15 AA.
XX
AC   ADR31534;
XX
DT   04-NOV-2004  (first entry)
XX
DE   IkappaB-alpha mutant substrate peptide #2.
XX
KW   Drug designing; clinical application; anti-inflammatory; anti-tumour;
KW   immune response; anti-ischaemic; artherosclerosis; therapy;
KW   IkappaB-alpha; mutant; mutein.
XX
```

OS Unidentified.
XX
FH Key Location/Qualifiers
FT Misc-difference 11
FT /note= "Wild type Ser substituted with Xaa"
FT /note= "Xaa corresponds to Asp, Glu"
XX
PN US2004157272-A1.
XX
PD 12-AUG-2004.
XX
PF 03-FEB-2004; 2004US-00771035.
XX
PR 09-MAR-2002; 2002US-00093945.
XX
PA (MERR-) MERRIMACK PHARM INC.
XX
PI Cardone MH, Yaffe M;
XX
DR WPI; 2004-580267/56.
XX
PT Identifying molecule capable of modulating kinase activity in situ, by
PT exposing candidate molecule to cell comprising signaling enzyme altered
PT to bind substrate having label and determining whether molecule changes
PT expression of label.
XX
PS Disclosure; Page 7; 25pp; English.
XX
CC The invention relates to a method for identifying molecule capable of
CC modulating kinase activity in situ. The method involves exposing
CC candidate molecule to cell having signalling enzyme altered to bind
CC phosphorylation substrate associated with detectable label where the
CC binding between altered signalling enzyme and substrate is regulated by a
CC kinase and determining if the candidate molecule cause a change in
CC expression of label. The method is useful for testing and designing drugs
CC with various clinical application e.g. anti-inflammatory candidate
CC molecules. The invention is useful in the field of anti-tumour
CC therapeutics and immune response regulating drugs. The method is also
CC useful for developing anti-ischaemic drugs that are useful for treating
CC artherosclerosis. The present sequence is IkappaB-alpha mutant substrate
CC peptide. This sequence is used to illustrate the method of the invention.
XX
SQ Sequence 15 AA;

Query Match 37.5%; Score 30; DB 8; Length 15;
Best Local Similarity 62.5%; Pred. No. 3e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 MRTRHSTG 12
:| || ||
Db 1 LRDRHDTG 8

RESULT 11
AAW59304
ID AAW59304 standard; peptide; 7 AA.
XX
AC AAW59304;
XX
DT 24-SEP-1998 (first entry)
XX
DE Non-polio enterovirus peptide fragment 11A.
XX
KW Non-polio enterovirus; NPEV; enteroviral disease; aseptic meningitis;
KW vaccination.
XX
OS Enterovirus sp.
XX
PN WO9814611-A2.
XX
PD 09-APR-1998.
XX
PF 01-OCT-1997; 97WO-US017734.
XX
PR 02-OCT-1996; 96US-0027353P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Kilpatrick D;
XX
DR WPI; 1998-240106/21.
XX
PT Identifying non-polio enteroviruses - using primers which hybridise to
PT sense and antisense strands that encode conserved non-polio enterovirus
PT peptide sequences.
XX
PS Claim 4; Page 23; 47pp; English.
XX
CC The peptide sequences AAW59298-W59344 are amplified by primers to detect
CC the presence of a non-polio enterovirus (NPEV)in a sample. The primers
CC and assays are used to detect NPEVs in a sample, to serotype these
CC viruses, to diagnose enteroviral diseases and medical conditions, and to
CC correlate (or disprove a correlation between) specific symptoms or
CC combinations of symptoms with the presence of a particular enterovirus.
CC They can be used for diseases such as aseptic meningitis. The detection
CC of NPEV infections and their correlation with medical conditions will
CC make possible vaccines and methods of treatment
XX
SQ Sequence 7 AA;

Query Match 36.2%; Score 29; DB 2; Length 7;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 TMRTRH 9
||:|||
Db 1 TMQTRH 6

RESULT 12
AAY50064
ID AAY50064 standard; peptide; 7 AA.
XX
AC AAY50064;
XX
DT 19-JAN-2000 (first entry)
XX
DE Coxsackievirus B VP1 conserved epitope 11.
XX
KW Virus; epitope; target; degenerate; PCR; primer; amplification; VP1;
KW nonstructural protein 2A; conserved; base analogue; inosine;
KW predetermined nucleotide; diagnosis; enterovirus; poliovirus.
XX
OS Synthetic.
OS Coxsackievirus.
XX
PN WO9953097-A2.
XX
PD 21-OCT-1999.
XX
PF 06-APR-1999; 99WO-US007513.
XX
PR 15-APR-1998; 98US-0081944P.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Kilpatrick DR;
XX
DR WPI; 1999-620444/53.
DR N-PSDB; AAZ30983.
XX
PT Designing degenerate polymerase chain reaction primers.
XX
PS Example 3; Page 18; 30pp; English.
XX
CC This sequence represents a conserved Coxsackievirus B (serotype B1) VP1

```
CC   epitope. The invention relates to a novel method for designing degenerate
CC   PCR primers (AAZ30975-Z31000, AAZ32601- AAZ32611) for amplifying target
CC   polynucleotides. This method comprises identifying uniquely conserved
CC   amino acid sequences (e.g, this epitope) in target proteins; synthesising
CC   degenerate polynucleotides encoding the conserved sequences; and
CC   substituting the synthesised polynucleotides with up to four
CC   predetermined nucleotides (e.g., inosine) at degenerate nucleotide
CC   positions. The nucleic acids comprise no more than 7 degenerate
CC   positions, have no more than 2 adjacent predetermined nucleotides and the
CC   predetermined nucleotides are 3 bases away from the 3' end of the
CC   synthesised strand. The degenerate primers are useful for amplifying
CC   target polynucleotides by the polymerase chain reaction (PCR). The use of
CC   the method of designing degenerate primers useful for the detection of
CC   polioviruses in clinical samples is described in US585477. The degenerate
CC   primers facilitate PCR amplification of unknown polynucleotides, where
CC   the amino acid sequence encoded is known. The primers also allow for the
CC   correlation of the subsequent molecular based diagnosis with a
CC   serologically derived diagnosis
XX
SQ   Sequence 7 AA;

  Query Match             36.2%;  Score 29;  DB 2;  Length 7;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 TMRTRH 9
              ||:|||
Db          1 TMQTRH 6


RESULT 13
AAB26509
ID   AAB26509 standard; peptide; 11 AA.
XX
AC   AAB26509;
XX
DT   11-JAN-2001  (first entry)
XX
DE   Human IgE C epsilon3 domain mimotope P11a.
XX
KW   IgE; C epsilon3; C epsilon4; histamine release inhibitor; vaccine;
KW   antibody; epitope; mimotope; human.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "N-terminal acetyl"
XX
PN   WO200050461-A1.
XX
PD   31-AUG-2000.
XX
PF   22-FEB-2000; 2000WO-EP001456.
XX
PR   25-FEB-1999;   99GB-00004408.
PR   21-JUL-1999;   99GB-00017144.
PR   07-AUG-1999;   99GB-00018598.
PR   07-AUG-1999;   99GB-00018599.
PR   07-AUG-1999;   99GB-00018601.
PR   07-AUG-1999;   99GB-00018604.
PR   07-AUG-1999;   99GB-00018606.
PR   29-OCT-1999;   99GB-00025618.
XX
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
PA   (PEPT-) PEPTIDE THERAPEUTICS LTD.
XX
PI   Friede M,  Mason S,  Turnell WG,  Van Mechelen MP;
PI   Vinals Y De BassolsC;
XX
DR   WPI; 2000-572074/53.
XX
PT   Peptides comprising surface exposed epitopes or mimotopes derived from C-
PT   epsilon-3 or C-epsilon-4 domains of IgE, useful for preventing or
PT   treating allergy.
XX
PS   Disclosure; Page 8; 76pp; English.
XX
CC   The present invention relates epitopes and mimotopes of an isolated
CC   surface exposed epitope of C epsilon3 or C epsilon4 domain of IgE. The
CC   epitopes were identified by calculating the accessible surface of each
CC   IgE residue. Mimotopes were designed to be similar to the epitopes. The
CC   epitopes are useful in preparing medicaments for treating or preventing
CC   allergies. The epitopes and mimotopes of the invention induce anti-IgE
CC   antibodies which are capable of raising non-anaphylactic antibodies and
CC   inhibiting histamine release. The present sequence is a IgE C epsilon3
CC   domain mimotope
XX
SQ   Sequence 11 AA;

  Query Match             36.2%;  Score 29;  DB 3;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 HSTGG 13
              |||||
Db          6 HSTGG 10


RESULT 14
AAU16840
ID   AAU16840 standard; peptide; 11 AA.
XX
AC   AAU16840;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Peptide P11a derived as mimotope of Cepsilon3/4 region of human IgE.
XX
KW   Human; linkage technology; conjugated compound; carrier vehicle; epitope;
KW   Cepsilon2; Cepsilon3; Cepsilon4; immunoglobulin E; IgE mediated disease;
KW   antibody response.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200145745-A2.
XX
PD   28-JUN-2001.
XX
PF   21-DEC-2000; 2000WO-GB004935.
XX
PR   21-DEC-1999;   99GB-00030233.
PR   22-FEB-2000; 2000GB-00004096.
PR   22-AUG-2000; 2000GB-00020707.
PR   22-AUG-2000; 2000GB-00020708.
XX
PA   (ACAM-) ACAMBIS RES LTD.
PA   (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI   Flinn N,  Johnson T;
XX
DR   WPI; 2001-521967/57.
XX
PT   A linkage comprising an immunogenic conjugate useful treatment of IgE
PT   mediated diseases.
XX
PS   Example 5; Page 25; 48pp; English.
XX
CC   The present invention relates to linkage methodology for use in the
CC   conjugation of compounds (e.g. peptides) to carrier vehicles (e.g.
CC   macromolecules, polymers, dendrimers, proteins) to produce biological and
CC   immunological constructs. The invention provides a method for linking an
CC   epitope (e.g. a peptide) to a carrier (e.g. a protein) for use in a
CC   pharmaceutical composition or a vaccine. The invention describes peptides
```

CC derived from or mimotopes of the Cepsilon2, Cepsilon3 or Cepsilon4
CC regions of human immunoglobulin E (IgE) which are used to produce
CC conjugated compounds. The compounds or compositions of the invention are
CC useful in the manufacture of a medicament for the treatment of IgE
CC mediated diseases. The invention allows for controlled conjugation of a
CC peptide epitope (antigen) to a protein so as to form an immunogenic
CC conjugate which may be able to raise a protective antibody response in an
CC animal or human patient. AAU16632-AAU16913 represent peptides derived
CC from or mimotopes of the Cepsilon2/Cepsilon3/Cepsilon4 region of human
CC IgE
XX
SQ Sequence 11 AA;

Query Match 36.2%; Score 29; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 HSTGG 13
|||||
Db 6 HSTGG 10

RESULT 15
ABJ00284
ID ABJ00284 standard; peptide; 11 AA.
XX
AC ABJ00284;
XX
DT 02-SEP-2002 (first entry)
XX
DE Human IgE immunogenic peptide SEQ ID NO: 68.
XX
KW Immunogen; human; IgE; immunoglobulin E; allergy; thio-ether linkage;
KW vaccine; antiallergic.
XX
OS Homo sapiens.
XX
PN WO200216409-A2.
XX
PD 28-FEB-2002.
XX
PF 17-AUG-2001; 2001WO-EP009576.
XX
PR 22-AUG-2000; 2000GB-00020717.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
PA (PEPT-) PEPTIDE THERAPEUTICS LTD.
XX
PI Friede M, Mason S, Turnell WG, Vinals Y BassolsC;
XX
DR WPI; 2002-489648/52.
XX
PT Conjugate for use in vaccine for treatment of allergy, comprises
PT disulfide bridge cyclized peptide and immunogenic carrier.
XX
PS Claim 4; Page 10; 45pp; English.
XX
CC The present invention relates to conjugates suitable for use in vaccines,
CC where the conjugate comprises a disulphide bridge cyclised peptide and an
CC immunogenic carrier. The vaccines can be used in the treatment of
CC allergies. The present sequence is a peptide immunogen derived from human
CC immunoglobulin E (IgE) suitable to be cyclised and used in the invention
XX
SQ Sequence 11 AA;

Query Match 36.2%; Score 29; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9 HSTGG 13
|||||
Db 6 HSTGG 10

Search completed: February 22, 2005, 09:24:42
Job time : 67.6667 secs

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-8
Perfect score: 76
Sequence: 1 FRQLVHFVRDFAQLL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	21	27.6	15	2	C48401	ribosomal protein
2	21	27.6	15	4	I52698	hypothetical THRA1
3	19	25.0	12	2	S68271	major urinary prot
4	19	25.0	14	2	C39170	acyl-[acyl-carrier
5	19	25.0	15	2	S29485	GTP-binding protei
6	18	23.7	7	2	B39127	phosphotransferase
7	18	23.7	10	2	A58365	neuropeptide FFRFa
8	18	23.7	10	2	S23307	neurokinin A - rai
9	18	23.7	10	2	S23186	neurokinin A - Atl
10	18	23.7	11	2	PD0441	translation elonga
11	18	23.7	12	2	S65730	hemoglobin, extrac
12	18	23.7	12	2	PT0228	Ig heavy chain CDR
13	18	23.7	14	2	PA0104	protein QF200070 -
14	18	23.7	14	2	S38307	DEB-A protein - fr
15	18	23.7	14	2	S29632	xylan 1,4-beta-xyl
16	18	23.7	14	2	A47421	leukotriene B-4 12
17	18	23.7	15	2	PQ0193	stylar glycoprotei
18	18	23.7	15	2	PQ0194	Sz-glycoprotein -
19	18	23.7	15	2	PA0054	protein QF200017 -
20	17	22.4	9	2	C36730	hutU protein - Kle
21	17	22.4	9	2	A44787	calliFMRFamide 10
22	17	22.4	10	2	PT0284	Ig heavy chain CRD
23	17	22.4	11	2	PH1600	Ig H chain V-D-J r
24	17	22.4	12	2	PH1182	T-cell receptor al
25	17	22.4	12	2	PH1174	T-cell receptor al
26	17	22.4	12	2	PH1181	T-cell receptor al
27	17	22.4	12	2	A34858	proteinase E - bla
28	17	22.4	15	2	S26525	T-cell receptor al
29	17	22.4	15	2	S26518	T-cell receptor al
30	17	22.4	15	2	S26517	T-cell receptor al
31	17	22.4	15	2	JN0730	hypothetical 1.7K
32	17	22.4	15	2	PH1443	T-cell receptor al
33	17	22.4	15	2	PH1441	T-cell receptor al
34	17	22.4	15	2	PQ0780	NADH2 dehydrogenas
35	16	21.1	9	2	PC2021	oxytocin-related p
36	16	21.1	10	2	B45482	platelet activatin
37	16	21.1	10	2	D37397	hypothetical prote
38	16	21.1	10	2	B46453	e antigen p20e pre
39	16	21.1	10	2	G58501	48K bile/gallbladd
40	16	21.1	11	2	S21127	precorrin methyltr
41	16	21.1	11	2	T06383	hypothetical prote
42	16	21.1	11	4	S41909	hypothetical prote
43	16	21.1	12	2	S68402	NAD(+)-glycohydrol
44	16	21.1	12	2	B56049	urinary tract ston
45	16	21.1	13	2	PS0453	36K protein 3124 -

ALIGNMENTS

RESULT 1
C48401
ribosomal protein L1 - Thermus aquaticus (fragment)
C;Species: Thermus aquaticus
C;Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: C48401
R;Garber, M.B.; Agalarov, S.C.; Eliseikina, I.A.; Fomenkova, N.P.; Nikonov, S.V.; Sedelni
Biochimie 74, 327-336, 1992
A;Title: Ribosomal proteins from Thermus thermophilus for structural investigations.
A;Reference number: A48401; MUID:92345325; PMID:1637860
A;Accession: C48401
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <GAR>
A;Note: sequence extracted from NCBI backbone (NCBIP:109932)

Query Match 27.6%; Score 21; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 FRQLVHFV 8
              :| |: ||
Db          7 YRALLXFV 14
```

RESULT 2
I52698
hypothetical THRA1/BTR mutant fusion protein, cell line BT474 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Apr-2000
C;Accession: I52698
R;Futreal, P.A.; Cochran, C.; Marks, J.R.; Iglehart, J.D.; Zimmerman, W.; Barrett, J.C.;
Cancer Res. 54, 1791-1794, 1994
A;Title: Mutation analysis of the THRA1 gene in breast cancer: deletion/fusion of the ge
A;Reference number: I52698; MUID:94185019; PMID:7511052
A;Accession: I52698
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-15 <FUT>
A;Cross-references: GB:S71020; NID:g546111; PIDN:AAB30341.1; PID:g546112
C;Comment: This sequence is the chimeric product of a deletion or translocation mutation
C;Genetics:
A;Gene: THRA1/BTR
A;Map position: 17q11.2
C;Keywords: fusion protein

Query Match 27.6%; Score 21; DB 4; Length 15;
Best Local Similarity 62.5%; Pred. No. 1.9e+03;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 FRQLVHFV 8
              ||  || |
```

```
Db          7 FRVQVHSV 14


RESULT 3
S68271
major urinary protein VIII - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S68271
R;Robertson, D.H.L.; Cox, K.A.; Gaskell, S.J.; Evershed, R.P.; Beynon, R.J.
Biochem. J. 316, 265-272, 1996
A;Title: Molecular heterogeneity in the major urinary proteins of the house mouse Mus mu
A;Reference number: S68271; MUID:96235201; PMID:8645216
A;Accession: S68271
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <ROB>

  Query Match             25.0%;  Score 19;  DB 2;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 FAQL 14
              ||||
Db          3 FAQL 6


RESULT 4
C39170
acyl-[acyl-carrier-protein] desaturase (EC 1.14.19.2) - avocado (fragments)
C;Species: Persea americana (avocado)
C;Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 03-Jun-2002
C;Accession: C39170
R;Shanklin, J.; Somerville, C.
Proc. Natl. Acad. Sci. U.S.A. 88, 2510-2514, 1991
A;Title: Stearoyl-acyl-carrier-protein desaturase from higher plants is structurally unr
A;Reference number: A39170; MUID:91172837; PMID:2006187
A;Accession: C39170
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <SHA>
C;Keywords: oxidoreductase

  Query Match             25.0%;  Score 19;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 3.9e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 DFAQLL 15
              |:| :|
Db          8 DYADIL 13


RESULT 5
S29485
GTP-binding protein o-ral - Pacific electric ray (fragment)
C;Species: Torpedo californica (Pacific electric ray)
C;Date: 22-Nov-1993 #sequence_revision 27-Feb-1997 #text_change 13-Mar-1997
C;Accession: S29485
R;Volknandt, W.; Pevsner, J.; Elferink, L.A.; Scheller, R.H.
FEBS Lett. 317, 53-56, 1993
A;Title: Association of three small GTP-binding proteins with cholinergic synaptic vesic
A;Reference number: S29485; MUID:93154521; PMID:8428634
A;Accession: S29485
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <VOL>

  Query Match             25.0%;  Score 19;  DB 2;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 4.2e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 FVRDF 11
              || |:

Db          5 FVEDY 9


RESULT 6
B39127
phosphotransferase system enzyme II (EC 2.7.1.69) - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 08-Oct-1999
C;Accession: B39127
R;Hardesty, C.; Ferran, C.; DiRienzo, J.M.
J. Bacteriol. 173, 449-456, 1991
A;Title: Plasmid-mediated sucrose metabolism in Escherichia coli: characterization of scr
rin.
A;Reference number: A39127; MUID:91100329; PMID:1846143
A;Accession: B39127
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-7 <HAR>
A;Cross-references: GB:M38416; NID:g155142; PIDN:AAA98418.1; PID:g155144
C;Keywords: phosphotransferase

  Query Match             23.7%;  Score 18;  DB 2;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 DFAQL 14
              || |:
Db          2 DFEQI 6


RESULT 7
A58365
neuropeptide FFRFamide - blue mussel
N;Alternate names: FMRFamide-related decapeptide; Mytilus FFRFamide
C;Species: Mytilus edulis (blue mussel)
C;Date: 20-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C;Accession: A58365
R;Fujisawa, Y.; Ikeda, T.; Nomoto, K.; Yasuda-Kamatani, Y.; Minakata, H.; Kenny, P.T.M.;
Comp. Biochem. Physiol. C 102, 91-95, 1992
A;Title: The FMRFamide-related decapeptide of Mytilus contains a D-amino acid residue.
A;Reference number: A58365; MUID:93047882; PMID:1358533
A;Accession: A58365
A;Molecule type: protein
A;Residues: 1-10 <FUJ>
A;Cross-references: UNIPROT:P42560
A;Experimental source: anterior byssus retractor muscle
C;Keywords: amidated carboxyl end; D-amino acid; neuropeptide
F;2/Modified site: D-leucine (Leu) #status experimental
F;10/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             23.7%;  Score 18;  DB 2;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 HFVR 9
              || |
Db          6 HFFR 9


RESULT 8
S23307
neurokinin A - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Aug-2004
C;Accession: S23307
R;Jensen, J.; Conlon, J.M.
Eur. J. Biochem. 206, 659-664, 1992
A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the cod
A;Reference number: S23186; MUID:92298992; PMID:1376687
A;Accession: S23307
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <JEN>
```

```
A;Cross-references: UNIPROT:P28500

  Query Match             23.7%;  Score 18;  DB 2;  Length 10;
  Best Local Similarity   30.0%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 HFVRDFAQLL 15
              | :  |  |:
Db          1 HKINSFVGLM 10


RESULT 9
S23186
neurokinin A - Atlantic cod
C;Species: Gadus morhua (Atlantic cod)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Aug-2004
C;Accession: S23186
R;Jensen, J.; Conlon, J.M.
Eur. J. Biochem. 206, 659-664, 1992
A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the cod
A;Reference number: S23186; MUID:92298992; PMID:1376687
A;Accession: S23186
A;Molecule type: protein
A;Residues: 1-10 <JEN>
A;Cross-references: UNIPROT:P28500
A;Experimental source: brain
C;Function:
A;Description: may play a physiological role in the regulation of cardiovascular and gas
A;Note: neurokinin A is derived by post-translational processing of preprotachykinin A
C;Keywords: neuropeptide; amidated carboxyl end; tachykinin
F;10/Modified site: amidated carboxyl end (Met) #status predicted

  Query Match             23.7%;  Score 18;  DB 2;  Length 10;
  Best Local Similarity   30.0%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          6 HFVRDFAQLL 15
              | :  |  |:
Db          1 HKINSFVGLM 10


RESULT 10
PD0441
translation elongation factor TU-like protein P43, mitochondrial - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 21-Aug-1998
C;Accession: PD0441
R;Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.
submitted to JIPID, August 1998
A;Description: Proteome analysis of mouse brain.
A;Reference number: PD0441
A;Accession: PD0441
A;Molecule type: protein
A;Residues: 1-11 <KAW>
A;Experimental source: striatum
C;Keywords: mitochondrion

  Query Match             23.7%;  Score 18;  DB 2;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 4.5e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 FVRD 10
              :|||
Db          8 YVRD 11


RESULT 11
S65730
hemoglobin, extracellular, component - earthworm (Lumbricus terrestris) (fragment)
C;Species: Lumbricus terrestris (common earthworm)
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C;Accession: S65730
R;Fushitani, K.; Higashiyama, K.; Asao, M.; Hosokawa, K.
Biochim. Biophys. Acta 1292, 273-280, 1996
A;Title: Characterization of the constituent polypeptides of the extracellular hemoglobi
A;Reference number: S65721; MUID:96176855; PMID:8597573
A;Accession: S65730
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <FUS>

  Query Match             23.7%;  Score 18;  DB 2;  Length 12;
  Best Local Similarity   40.0%;  Pred. No. 5e+03;
  Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 HFVRD 10
              | ::|
Db          8 HLIQD 12


RESULT 12
PT0228
Ig heavy chain CDR3 region (clone 1-112) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0228
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and jo
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0228
A;Molecule type: DNA
A;Residues: 1-12 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

  Query Match             23.7%;  Score 18;  DB 2;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 5e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 FVRD 10
              :|||
Db          3 YVRD 6


RESULT 13
PA0104
protein QF200070 - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 23-Mar-2001
C;Accession: PA0104
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrichi
A;Reference number: PA0051
A;Accession: PA0104
A;Molecule type: protein
A;Residues: 1-14 <CHO>

  Query Match             23.7%;  Score 18;  DB 2;  Length 14;
  Best Local Similarity   60.0%;  Pred. No. 5.8e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RQLVH 6
              ||:|:
Db         10 RQIVY 14


RESULT 14
S38307
DEB-A protein - fruit fly (Drosophila melanogaster) (fragment)
C;Species: Drosophila melanogaster
C;Date: 19-May-1994 #sequence_revision 27-Feb-1997 #text_change 17-Mar-1999
C;Accession: S38307
R;Wang, G.L.; Goldstein, E.S.
Biochim. Biophys. Acta 1216, 94-104, 1993
```

A;Title: An AP-1 binding site in the upstream region of DEB-A is part of a developmental
A;Reference number: S38307; MUID:94032494; PMID:8218421
A;Accession: S38307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-14 <WAN>

Query Match 23.7%; Score 18; DB 2; Length 14;
Best Local Similarity 33.3%; Pred. No. 5.8e+03;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          7 FVRDFA 12
              ::| |:
Db          6 YIRSFS 11
```

RESULT 15
S29632
xylan 1,4-beta-xylosidase (EC 3.2.1.37) - Thermotoga sp. (strain FjSS3-B.1) (fragment)
N;Alternate names: beta-xylosidase
C;Species: Thermotoga sp.
A;Variety: FjSS3-B.1
C;Date: 19-Mar-1997 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: S29632
R;Ruttersmith, L.D.; Daniel, R.M.
Biochim. Biophys. Acta 1156, 167-172, 1993
A;Title: Thermostable beta-glucosidase and beta-xylosidase from Thermotoga sp. strain Fj
A;Reference number: S29631; MUID:93152594; PMID:8427876
A;Accession: S29632
A;Molecule type: protein
A;Residues: 1-14 <RUT>
A;Cross-references: UNIPROT:Q7M0Q6
A;Experimental source: strain FjSS3-B.1
C;Comment: Although the beta-xylosidase enzyme activity was apparently confirmed for thi
C;Function:
A;Description: hydrolyzes short chain oligosacchrides and xylobiose to produce D-xylose
A;Note: plays an important role in the relief of end-product inhibition of endoxylanase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 23.7%; Score 18; DB 2; Length 14;
Best Local Similarity 33.3%; Pred. No. 5.8e+03;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 RQLVHFVRD 10
              :: |:|  |
Db          2 KKYVYFFAD 10
```

Search completed: February 22, 2005, 09:46:27
Job time : 12.1333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-8
Perfect score: 76
Sequence: 1 FRQLVHFVRDFAQLL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	34.2	15	2	Q54325	Q54325 staphylococ
2	23	30.3	9	2	Q8GIZ6	Q8giz6 lactobacill
3	23	30.3	13	2	Q7XB02	Q7xb02 zea mays (m
4	23	30.3	13	2	Q6XFQ8	Q6xfq8 bacillus cl
5	23	30.3	14	2	Q7X8P6	Q7x8p6 zea mays (m
6	23	30.3	15	2	Q7XB01	Q7xb01 zea mays (m
7	22	28.9	12	2	Q9U7R8	Q9u7r8 conus querc
8	21	27.6	10	2	Q9TS43	Q9ts43 sus scrofa
9	21	27.6	11	2	Q90735	Q90735 gallus gall
10	21	27.6	13	2	Q6LAU9	Q6lau9 homo sapien
11	20	26.3	10	2	Q9TWU1	Q9twu1 fusinus fer
12	20	26.3	10	2	Q6R7V4	Q6r7v4 carlia zuma
13	20	26.3	11	2	Q68LE0	Q68le0 pyriglena l
14	20	26.3	11	2	Q68LE9	Q68le9 myrmotherul
15	20	26.3	13	2	Q6TKD3	Q6tkd3 praecitrull
16	20	26.3	13	2	Q6TKD4	Q6tkd4 sechium edu
17	20	26.3	13	2	Q6TKD5	Q6tkd5 sicyos angu
18	20	26.3	13	2	Q6TKD6	Q6tkd6 trichosanth
19	20	26.3	13	2	Q6TKD7	Q6tkd7 luffa grave
20	20	26.3	13	2	Q6TKD8	Q6tkd8 luffa echin
21	20	26.3	13	2	Q6TKD9	Q6tkd9 cucurbita p
22	20	26.3	13	2	Q6TKE0	Q6tke0 benincasa h
23	20	26.3	13	2	Q6TKE1	Q6tke1 marah orega
24	20	26.3	13	2	Q6TKE2	Q6tke2 cyclanthera
25	20	26.3	13	2	Q6TKE3	Q6tke3 echinocysti
26	20	26.3	13	2	Q6TKE4	Q6tke4 luffa quinq
27	20	26.3	13	2	Q6TKE5	Q6tke5 lagenaria l
28	20	26.3	13	2	Q6TKE6	Q6tke6 citrullus l
29	20	26.3	13	2	Q6TKE7	Q6tke7 citrullus c
30	20	26.3	13	2	Q6TKE8	Q6tke8 acanthosicy
31	20	26.3	13	2	Q6TKE9	Q6tke9 bryonia dio
32	20	26.3	13	2	Q6TKF0	Q6tkf0 coccinia pa
33	20	26.3	13	2	Q6TKF1	Q6tkf1 diplocyclos
34	20	26.3	13	2	Q6TKF2	Q6tkf2 ecballium e
35	20	26.3	13	2	Q9MQK3	Q9mqk3 capra ibex
36	20	26.3	14	2	Q14342	Q14342 homo sapien
37	20	26.3	14	2	Q6R7V0	Q6r7v0 carlia viva
38	20	26.3	15	2	Q6LAI1	Q6lai1 drosophila
39	19	25.0	10	1	AKHX_LOCMI	P81626 locusta mig
40	19	25.0	11	2	Q8MM58	Q8mm58 heliconius
41	19	25.0	11	2	Q9QYF6	Q9qyf6 mus musculu
42	19	25.0	12	2	Q8MUN4	Q8mun4 heliconius
43	19	25.0	12	2	Q8MUN9	Q8mun9 heliconius
44	19	25.0	12	2	Q9T1J4	Q9t1j4 bacteriopha
45	19	25.0	13	1	TEML_RANTE	P57104 rana tempor

ALIGNMENTS

RESULT 1
Q54325
ID Q54325 PRELIMINARY; PRT; 15 AA.
AC Q54325;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Chloramphenicol acetyltransferase (Fragment).
GN Name=cat;
OS Staphylococcus intermedius.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1285;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96379895; PubMed=8787908;
RA Lodder G., Schwarz S., Gregory P., Dyke K.;
RT "Tandem duplication in ermC translational attenuator of the macrolide-
RT lincosamide-streptogramin B resistance plasmid pSES6 from
RT Staphylococcus equorum.";
RL Antimicrob. Agents Chemother. 40:215-217(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Schwarz S.P.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.
DR EMBL; X82666; CAA57981.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Transferase.
FT NON_TER 1 1
SQ SEQUENCE 15 AA; 1931 MW; 173A6AA93999A8D9 CRC64;

Query Match 34.2%; Score 26; DB 2; Length 15;
Best Local Similarity 36.4%; Pred. No. 1.4e+03;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 FRQLVHFVRDF 11
              |: ::| | |:
Db          4 FQDIIHRVDDW 14
```

RESULT 2
Q8GIZ6
ID Q8GIZ6 PRELIMINARY; PRT; 9 AA.
AC Q8GIZ6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glyceraldehyde 3-phosphate dehydrogenase (EC 1.2.1.12)
DE (Fragment).
GN Name=gap;
OS Lactobacillus delbrueckii (subsp. lactis).
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=29397;
RN [1]

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC88;
RA   Bourniquel A.A., Mollet B.;
RT   "Purification and characterization of the 3-phosphoglycerate kinase
RT   from the thermophile Lactobacillus delbrueckii subsp. lactis.";
RL   Int. Dairy J. 12:723-728(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC88;
RA   Bourniquel A.A.;
RT   "Molecular insights into the metabolism and physiology of the lactic
RT   acid bacterium Lactobacillus delbrueckii subsp. lactis.";
RL   Thesis (2000), Department of Molecular Microbiology, Biozentrum der
RL   Universitaet Basel, (PhD work conducted at the Nestle Research Center,
RL   Lausanne), Switzerland.
DR   EMBL; AJ515554; CAD56494.1; -.
DR   GO; GO:0004365; F:glyceraldehyde-3-phosphate dehydrogenase (p. . .; IEA.
DR   GO; GO:0016491; F:oxidoreductase activity; IEA.
KW   Oxidoreductase.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1071 MW;  94ABADD9C1E72731 CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RQLVHF 7
              | |:||
Db          1 RTLLHF 6


RESULT 3
Q7XB02
ID   Q7XB02      PRELIMINARY;      PRT;    13 AA.
AC   Q7XB02;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Phytoene synthase 2 (Fragment).
GN   Name=psy2;
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B73;
RX   MEDLINE=22779048; PubMed=12897253;
RA   Palaisa K.A., Morgante M., Williams M., Rafalski A.;
RT   "Contrasting effects of selection on sequence diversity and linkage
RT   disequilibrium at two phytoene synthase loci.";
RL   Plant Cell 15:1795-1806(2003).
DR   EMBL; AY300613; AAP55352.1; -.
FT   NON_TER       1      1
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1443 MW;  1D425AC312022054 CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 13;
  Best Local Similarity   36.4%;  Pred. No. 4.1e+03;
  Matches    4;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 QLVHFVRDFAQ 13
              || : :||  :
Db          3 QLTNILRDVGE 13


RESULT 4
Q6XFQ8
ID   Q6XFQ8      PRELIMINARY;      PRT;    13 AA.
AC   Q6XFQ8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Putative leader peptide.
OS   Bacillus clausii.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=79880;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM8716;
RX   PubMed=14711653; DOI=10.1128/AEM.70.1.280-284.2004;
RA   Bozdogan B., Galopin S., Leclercq R.;
RT   "Characterization of a new erm-related macrolide resistance gene
RT   present in probiotic strains of Bacillus clausii.";
RL   Appl. Environ. Microbiol. 70:280-284(2004).
DR   EMBL; AY234334; AAP74656.1; -.
SQ   SEQUENCE   13 AA;  1687 MW;  48F43CB7F010D407 CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 13;
  Best Local Similarity   60.0%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VHFVR 9
              :||:|
Db          1 MHFIR 5


RESULT 5
Q7X8P6
ID   Q7X8P6      PRELIMINARY;      PRT;    14 AA.
AC   Q7X8P6;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Phytoene synthase 2 (Fragment).
GN   Name=psy2;
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PI583352, PI595566, Ames22026, and PI595546;
RX   MEDLINE=22779048; PubMed=12897253;
RA   Palaisa K.A., Morgante M., Williams M., Rafalski A.;
RT   "Contrasting effects of selection on sequence diversity and linkage
RT   disequilibrium at two phytoene synthase loci.";
RL   Plant Cell 15:1795-1806(2003).
DR   EMBL; AY300630; AAP55368.1; -.
DR   EMBL; AY300637; AAP55375.1; -.
DR   EMBL; AY300643; AAP55381.1; -.
DR   EMBL; AY300662; AAP55400.1; -.
DR   InterPro; IPR008949; Terpenoid_synth.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1556 MW;  1D425AC312580D54 CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   36.4%;  Pred. No. 4.4e+03;
  Matches    4;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 QLVHFVRDFAQ 13
              || : :||  :
Db          4 QLTNILRDVGE 14


RESULT 6
Q7XB01
ID   Q7XB01      PRELIMINARY;      PRT;    15 AA.
AC   Q7XB01;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Phytoene synthase 2 (Fragment).
```

```
GN   Name=psy2;
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PI270297;
RX   MEDLINE=22779048; PubMed=12897253;
RA   Palaisa K.A., Morgante M., Williams M., Rafalski A.;
RT   "Contrasting effects of selection on sequence diversity and linkage
RT   disequilibrium at two phytoene synthase loci.";
RL   Plant Cell 15:1795-1806(2003).
DR   EMBL; AY300625; AAP55363.1; -.
DR   InterPro; IPR008949; Terpenoid_synth.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1613 MW;  1D425AC312585ACF CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   36.4%;  Pred. No. 4.8e+03;
  Matches    4;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 QLVHFVRDFAQ 13
              || : :||   :
Db          5 QLTNILRDVGE 15


RESULT 7
Q9U7R8
ID   Q9U7R8      PRELIMINARY;      PRT;    12 AA.
AC   Q9U7R8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Calmodulin (Fragment).
OS   Conus quercinus (Oak cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
OX   NCBI_TaxID=101313;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99398698; PubMed=10468598; DOI=10.1073/pnas.96.18.10272;
RA   Duda T.F. Jr., Palumbi S.R.;
RT   "Developmental shifts and species selection in gastropods.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:10272-10277(1999).
DR   EMBL; AF113305; AAD49168.1; -.
FT   NON_TER       1      1
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1284 MW;  1CF4EE6A86D9CAA0 CRC64;

  Query Match             28.9%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 5.7e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 DFAQLL 15
              |||: |
Db          6 DFAEFL 11


RESULT 8
Q9TS43
ID   Q9TS43      PRELIMINARY;      PRT;    10 AA.
AC   Q9TS43;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   OESTRADIOL-RECEPTOR=P1 peptide (Fragment).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91291128; PubMed=2064608;
RA   Thole H.H., Jungblut P.W., Jakob F.;
RT   "The proton-driven dissociation of oestradiol-receptor dimers as a
RT   preparative tool. Isolation of a 32 kDa fragment from porcine uteri
RT   and assignment of C-terminal origin by partial sequencing.";
RL   Biochem. J. 276:709-714(1991).
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1240 MW;  D314D274405691F2 CRC64;

  Query Match             27.6%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 7.1e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 QLVHFV 8
              :||| :
Db          1 ELVHMI 6


RESULT 9
Q90735
ID   Q90735      PRELIMINARY;      PRT;    11 AA.
AC   Q90735;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Gallus gallus beta-globin gene. (Fragment).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=81208060; PubMed=6263308;
RA   Day L.E., Hirst A.J., Lai E.C., Mace M.Jr., Woo S.L.C.;
RT   "5' Domain and nucleotide sequence of an adult chicken chromosomal
RT   beta-globin gene.";
RL   Biochemistry 20:2091-2098(1981).
DR   EMBL; V00378; CAA23677.1; -.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1372 MW;  271C02021B1DC1B3 CRC64;

  Query Match             27.6%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   36.4%;  Pred. No. 7.9e+03;
  Matches    4;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 LVHFVRDFAQL 14
              :||:  :  ||
Db          1 MVHWTAEEKQL 11


RESULT 10
Q6LAU9
ID   Q6LAU9      PRELIMINARY;      PRT;    13 AA.
AC   Q6LAU9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   B2-bradykinin receptor gene protein (Fragment).
GN   Name=B2-bradykinin receptor gene;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95298027; PubMed=7779089;
RA   Kammerer S., Braun A., Arnold N., Roscher A.A.;
RT   "The human bradykinin B2 receptor gene: full length cDNA, genomic
```

```
RT   organization and identification of the regulatory region.";
RL   Biochem. Biophys. Res. Commun. 211:226-233(1995).
DR   EMBL; X86163; CAA60107.1; -.
DR   GO; GO:0004872; F:receptor activity; IEA.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1567 MW;  DDE0E98A14967AB6 CRC64;

  Query Match             27.6%;  Score 21;  DB 2;  Length 13;
  Best Local Similarity   37.5%;  Pred. No. 9.3e+03;
  Matches    3;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 VHFVRDFA 12
              :| ::|:|
Db          2 IHKLQDWA 9


RESULT 11
Q9TWU1
ID   Q9TWU1      PRELIMINARY;      PRT;    10 AA.
AC   Q9TWU1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   FMRFAMIDE-related peptide.
OS   Fusinus ferrugineus (Ferruginous spindle).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Buccinoidea; Fasciolariidae; Fusinus.
OX   NCBI_TaxID=6488;
RN   [1]
RP   SEQUENCE.
RA   Kuroki Y., Kanda T., Kubota I., Ikeda T., Fujisawa Y., Minakata H.,
RA   Muneoka Y.;
RT   "FMRFamide-related peptides isolated from the prosobranch mollusc
RT   Fusinus ferrugineus.";
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
SQ   SEQUENCE   10 AA;  1233 MW;  83F80229C1EAA451 CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 1.1e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 HFVR 9
              ||:|
Db          6 HFLR 9


RESULT 12
Q6R7V4
ID   Q6R7V4      PRELIMINARY;      PRT;    10 AA.
AC   Q6R7V4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN   Name=GAPDH;
OS   Carlia zuma.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Scincoidea;
OC   Scincidae; Carlia.
OX   NCBI_TaxID=260893;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dolman G., Phillips B.;
RT   "Single copy nuclear DNA markers characterized for comparative
RT   phylogeography in Australian wet tropics rainforest skinks.";
RL   Mol. Ecol. Notes 4:185-187(2004).
DR   EMBL; AY508912; AAS09890.1; -.
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1171 MW;  9D0ABB2322C9C1EA CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 1.1e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 HFVR 9
              |||:
Db          4 HFVK 7


RESULT 13
Q68LE0
ID   Q68LE0      PRELIMINARY;      PRT;    11 AA.
AC   Q68LE0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN   Name=NAPDH;
OS   Pyriglena leuconota.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Formicariidae;
OC   Pyriglena.
OX   NCBI_TaxID=183187;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15283860;
RA   Irestedt M., Fjeldsa J., Nylander J.A., Ericson P.G.;
RT   "Phylogenetic relationships of typical antbirds (Thamnophilidae) and
RT   test of incongruence based on Bayes factors.";
RL   BMC Evol. Biol. 4:23-23(2004).
DR   EMBL; AY677056; AAT96981.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1242 MW;  9D0AEB2622C9C1EA CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 1.2e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 HFVR 9
              |||:
Db          5 HFVK 8


RESULT 14
Q68LE9
ID   Q68LE9      PRELIMINARY;      PRT;    11 AA.
AC   Q68LE9;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN   Name=NAPDH;
OS   Myrmotherula fulviventris.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Thamnophilidae;
OC   Myrmotherula.
OX   NCBI_TaxID=288045;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15283860;
RA   Irestedt M., Fjeldsa J., Nylander J.A., Ericson P.G.;
RT   "Phylogenetic relationships of typical antbirds (Thamnophilidae) and
RT   test of incongruence based on Bayes factors.";
RL   BMC Evol. Biol. 4:23-23(2004).
DR   EMBL; AY677047; AAT96972.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1242 MW;  9D0AEB2622C9C1EA CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 1.2e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          6 HFVR 9
              |||:
Db          5 HFVK 8


RESULT 15
Q6TKD3
ID   Q6TKD3      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKD3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Praecitrullus fistulosus.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Praecitrullus.
OX   NCBI_TaxID=252558;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396191; AAR07576.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             26.3%;  Score 20;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 RDFAQLL 15
              |||  :|
Db          1 RDFEPVL 7


Search completed: February 22, 2005, 09:37:57
Job time : 53.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-8
Perfect score: 76
Sequence: 1 FRQLVHFVRDFAQLL 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	76	100.0	15	2	AAW04848	Aaw04848 Internal
2	35	46.1	15	3	AAB08960	Aab08960 Human sec
3	33	43.4	12	8	ADO06895	Ado06895 Porcine r
4	31	40.8	14	8	ADG98090	Adg98090 Apoptosis
5	31	40.8	15	6	ABP71311	Abp71311 Drosophil
6	31	40.8	15	6	ADA10648	Ada10648 Drosophil
7	30	39.5	14	2	AAY06719	Aay06719 Peptide f
8	29	38.2	9	7	ADC21478	Adc21478 Human PRD
9	29	38.2	10	7	ADC21483	Adc21483 Human PRD
10	28	36.8	9	4	AAG84689	Aag84689 MAGE3 cro
11	28	36.8	9	4	AAG84688	Aag84688 MAGE3 cro
12	28	36.8	9	4	AAU27075	Aau27075 Human Leu
13	28	36.8	9	4	AAU27076	Aau27076 Human Leu
14	28	36.8	9	4	AAU26742	Aau26742 Human Leu
15	28	36.8	9	4	AAU26743	Aau26743 Human Leu
16	28	36.8	9	4	AAB75861	Aab75861 Tumour as
17	28	36.8	9	8	ADP26069	Adp26069 Plasmodiu
18	28	36.8	11	3	AAB12656	Aab12656 Protein k
19	28	36.8	15	4	AAM99415	Aam99415 Vaccine r
20	28	36.8	15	4	AAU69222	Aau69222 Human tra
21	28	36.8	15	8	ADK69557	Adk69557 Epitope l
22	27	35.5	6	2	AAY24297	Aay24297 Somatosta
23	27	35.5	8	7	ADD57458	Add57458 HLA bindi
24	27	35.5	8	7	ADD56995	Add56995 HLA bindi
25	27	35.5	8	7	ADD57770	Add57770 HLA bindi
26	27	35.5	10	4	AAM42696	Aam42696 Mycoplasm
27	27	35.5	10	6	ABJ67822	Abj67822 184P1E2-r
28	27	35.5	10	6	ABJ69045	Abj69045 184P1E2-r
29	27	35.5	10	6	ABJ68535	Abj68535 184P1E2-r
30	27	35.5	10	6	ABJ66708	Abj66708 184P1E2-r
31	27	35.5	10	6	ABJ66180	Abj66180 184P1E2-r
32	27	35.5	10	6	ABJ67345	Abj67345 184P1E2-r
33	27	35.5	10	7	ADI20156	Adi20156 Anchor mo
34	27	35.5	10	7	ADI20288	Adi20288 Anchor mo
35	27	35.5	10	7	ADI20467	Adi20467 Anchor mo
36	27	35.5	10	7	ADI20298	Adi20298 Anchor mo
37	27	35.5	10	8	ADI53583	Adi53583 Non oncog
38	27	35.5	10	8	ADM33674	Adm33674 HPV E6 PD
39	27	35.5	10	8	ADR82734	Adr82734 Human pap
40	27	35.5	10	8	ADR82804	Adr82804 Human pap
41	27	35.5	12	2	AAY21356	Aay21356 Human HUP
42	27	35.5	14	2	AAW64250	Aaw64250 Mouse mas
43	27	35.5	15	2	AAW04851	Aaw04851 Internal
44	27	35.5	15	5	ABP63564	Abp63564 Human pap
45	27	35.5	15	6	ABR37049	Abr37049 Human can

ALIGNMENTS

RESULT 1
AAW04848
ID AAW04848 standard; peptide; 15 AA.
XX
AC AAW04848;
XX
DT 18-FEB-1997 (first entry)
XX
DE Internal fragment of herpes simplex virus UL15 protein.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Herpes simplex virus.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 2; Page 42; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is an internal peptide of herpes

CC simplex virus UL15 protein and is implicated as a foreign epitope
CC involved in the aetiology or in remissions of multiple sclerosis. It has
CC been shown capable of inducing the proliferation of autoreactive T-cell
CC clones isolated from multiple sclerosis patients
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FRQLVHFVRDFAQLL 15
||||||||||||||||
Db 1 FRQLVHFVRDFAQLL 15

RESULT 2
AAB08960
ID AAB08960 standard; protein; 15 AA.
XX
AC AAB08960;
XX
DT 30-AUG-2000 (first entry)
XX
DE Human secreted protein sequence encoded by gene 24 SEQ ID NO:117.
XX
KW Human; secreted protein; cytostatic; anti-proliferative; vulnerary;
KW immunosuppressive; antibacterial; diagnosis; immune system; chemotaxis;
KW hyperproliferative disorder; infectious disease; tissue regeneration;
KW screening; food additive; preservative; wound healing;
KW hyper-vascular disease.
XX
OS Homo sapiens.
XX
PN WO200017222-A1.
XX
PD 30-MAR-2000.
XX
PF 22-SEP-1999; 99WO-US022012.
XX
PR 23-SEP-1998; 98US-0101546P.
PR 02-OCT-1998; 98US-0102895P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Rosen CA, Duan RD, Shi Y, Lafleur DW, Young PE, Ni J;
PI Komatsoulis G, Endress GA, Soppet DR;
XX
DR WPI; 2000-283538/24.
XX
PT Human secreted proteins and coding sequences useful in diagnostic and
PT therapeutic methods for disorders such as immune system or proliferative
PT disorders, related to the proteins.
XX
PS Disclosure; Page 69; 416pp; English.
XX
CC The polynucleotide sequences given in AAA39052 to AAA39088 encode the
CC human secreted proteins given in AAB08891 to AAB08984. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic; anti-
CC proliferative; immunosuppressive; antibacterial; and vulnerary. The
CC secreted proteins and their related polynucleotide sequences are useful
CC for diagnostic and therapeutic methods useful for diagnosing and treating
CC disorders related to the secreted proteins. The proteins, and
CC polynucleotide sequences may be useful for treating disorders of the
CC immune system, hyperproliferative disorders, infectious disease,
CC regeneration of tissues, for chemotaxis and for screening molecules that
CC bind to the proteins. The proteins or polynucleotide sequences may be
CC used as food additives or preservatives, to increase or decrease storage
CC capabilities, fat content, lipid, protein, carbohydrate, vitamins,
CC minerals, co-factors or other nutritional components. Agonists or
CC antagonists of the proteins may be used to prevent scar tissue growth
CC during wound healing, and hyper-vascular diseases. AAA39043 to AAA39051
CC and AAB08890 are sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 15 AA;

Query Match 46.1%; Score 35; DB 3; Length 15;
Best Local Similarity 50.0%; Pred. No. 47;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LVHFVRDFAQLL 15
| ||: |:|:|:
Db 3 LFHFLIDYAELV 14

RESULT 3
ADO06895
ID ADO06895 standard; peptide; 12 AA.
XX
AC ADO06895;
XX
DT 01-JUL-2004 (first entry)
XX
DE Porcine reproductive and respiratory syndrome virus gp5 peptide #14.
XX
KW virucide; gene therapy; vaccine; neutralising epitope; mimotopes;
KW Porcine reproductive and respiratory syndrome virus; PRRSV; epitope B;
KW gp5.
XX
OS Porcine reproductive and respiratory syndrome virus.
XX
PN US2004014028-A1.
XX
PD 22-JAN-2004.
XX
PF 19-JUL-2002; 2002US-00199545.
XX
PR 19-JUL-2002; 2002US-00199545.
XX
PA (WOLF-) WOLF BIOTECH.
XX
PI Lopez OJ, Ostrowski M;
XX
DR WPI; 2004-121549/12.
XX
PT Isolating neutralizing epitopes of a pathogen for vaccination purposes
PT comprises obtaining sera with and without neutralizing antibodies from a
PT species and using the difference between the sera to isolate neutralizing
PT epitopes.
XX
PS Example 2; Fig 3; 28pp; English.
XX
CC The invention describes isolating neutralising epitopes of a pathogen
CC comprising obtaining sera with and without neutralising antibodies from a
CC species, and using the difference between the serum with neutralising
CC antibodies and the serum without neutralising antibodies to isolate
CC neutralising epitopes and/or mimotopes. The composition and methods are
CC useful in diagnosing, preventing or treating Porcine reproductive and
CC respiratory syndrome virus (PRRSV) infection. The polyclonal or
CC monoclonal antibodies specific to epitope B, or the monospecific chimeric
CC mouse/porcine antibodies obtained by recombinant DNA technology specific
CC to epitope B, are used to facilitate cure of infected animals or
CC prevention of animals at risk of infection with PRRSV. This is the amino
CC acid sequence of a porcine reproductive and respiratory syndrome virus
CC (PRRSV) gp5 peptide screened using anti-PRRSV specific swine-antibodies
CC to identify protective epitopes.
XX
SQ Sequence 12 AA;

Query Match 43.4%; Score 33; DB 8; Length 12;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 VHFVRDFAQ 13

```
              ||| | |:|
Db          4 VHFQRQFSQ 12


RESULT 4
ADG98090
ID   ADG98090 standard; peptide; 14 AA.
XX
AC   ADG98090;
XX
DT   11-MAR-2004  (first entry)
XX
DE   Apoptosis induction-related Drosophila melanogaster IAPi peptide SeqID3.
XX
KW   inhibitor of apoptosis protein; IAP; anti-HIV; nootropic;
KW   neuroprotective; vasotropic; cytostatic; immunosuppressive; caspase;
KW   inducing apoptosis; AIDS; IAPi; neurodegenerative disease;
KW   ischaemic injury; cancer; autoimmune disease; fruit fly.
XX
OS   Drosophila melanogaster.
XX
PN   WO2003095473-A2.
XX
PD   20-NOV-2003.
XX
PF   08-MAY-2003; 2003WO-US012125.
XX
PR   09-MAY-2002; 2002US-0378668P.
PR   24-FEB-2003; 2003US-0448869P.
XX
PA   (ROCK-) ROCKEFELLER INST.
PA   (RAPA-) RAPAPORT FAMILY INST RES IN MEDICAL SCI.
XX
PI   Steller H,  Don Ryoo H,  Ciechanover A,  Gonen H;
XX
DR   WPI; 2004-012082/01.
DR   N-PSDB; ADG98101.
XX
PT   New nucleic acid molecules encoding a peptide or polypeptide that binds
PT   to a portion of an inhibitor of apoptosis protein, useful for inducing
PT   apoptosis and identifying inhibitors or enhancers of apoptosis for
PT   treating AIDS, or cancer.
XX
PS   Claim 46; SEQ ID NO 3; 101pp; English.
XX
CC   This invention relates to a novel isolated nucleic acid molecule which
CC   encodes a polypeptide that specifically binds to at least a portion of an
CC   inhibitor of apoptosis protein (IAP). The invention may be useful for the
CC   development of compounds with an anti-HIV, nootropic, neuroprotective,
CC   vasotropic, cytostatic or immunosuppressive activity through the
CC   inhibition or stimulation of caspase. The nucleic acid molecules and
CC   peptides or polypeptides are useful for inducing apoptosis and
CC   identifying inhibitors or enhancers of apoptosis for treating AIDS,
CC   neurodegenerative diseases, ischaemic injury, cancer or autoimmune
CC   diseases. The present sequence is that of a fruit fly (D melanogaster)-
CC   derived IAPi (IAP inhibitor) peptide of the invention.
XX
SQ   Sequence 14 AA;

  Query Match             40.8%;  Score 31;  DB 8;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 2.1e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 LVHFVRDFAQLL 15
              : :|: | ||||
Db          2 IAYFIPDQAQLL 13


RESULT 5
ABP71311
ID   ABP71311 standard; peptide; 15 AA.
XX
AC   ABP71311;
XX
DT   28-APR-2003  (first entry)
XX
DE   Drosophila grim protein N-terminal fragment.
XX
KW   Omi; HtrA2; serine protease; inhibitor of apoptosis protein; IAP;
KW   caspase; apoptosis; cytostatic; immunosuppressive; neuroprotective;
KW   vasotropic; gene therapy; grim.
XX
OS   Drosophila sp.
XX
PN   WO2003006680-A2.
XX
PD   23-JAN-2003.
XX
PF   15-JUL-2002; 2002WO-US022658.
XX
PR   13-JUL-2001; 2001US-0305378P.
PR   14-DEC-2001; 2001US-0340163P.
XX
PA   (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI   Alnemri ES;
XX
DR   WPI; 2003-221760/21.
XX
PT   New Omi nucleic acids and peptides that bind to an inhibitor of apoptosis
PT   proteins, useful for regulating or altering caspase-mediated apoptosis
PT   and for treating cancer, tumor, or autoimmune diseases.
XX
PS   Example 2; Fig 6; 83pp; English.
XX
CC   The invention relates to polynucleotides encoding an Omi (serine
CC   protease) peptide or polypeptide. The Omi peptide specifically binds to a
CC   portion of an Inhibitor of Apoptosis Protein (IAP). The Omi polypeptide
CC   induces caspase-independent apoptosis, or fails to have serine protease
CC   activity. The Omi peptides are useful for regulating or altering
CC   apoptosis, specifically caspase-mediated apoptosis, and as immunogens for
CC   raising antibodies. Enhancers of apoptosis are useful for treating
CC   cancers, tumours or for destroying cells that mediate autoimmune
CC   diseases. Compositions may also be used for the treatment of diseases
CC   associated with inappropriate activation of apoptosis such as
CC   neurodegenerative diseases and ischaemic injury. The antibodies can be
CC   used in isolating Omi peptides, polypeptides and their variants, in
CC   identifying molecules that interact with Omi peptides and polypeptides,
CC   and in inhibiting or enhancing the biological activity of Omi peptides
CC   and polypeptides. Sequences ABP71310-315 represent fragments of various
CC   IAP-binding proteins, used to determine Omi as a IAP-binding protein
XX
SQ   Sequence 15 AA;

  Query Match             40.8%;  Score 31;  DB 6;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 2.3e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 LVHFVRDFAQLL 15
              : :|: | ||||
Db          2 IAYFIPDQAQLL 13


RESULT 6
ADA10648
ID   ADA10648 standard; peptide; 15 AA.
XX
AC   ADA10648;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Drosophila Grim protein N-terminus #1.
XX
KW   caspase-9; anti-HIV; nootropic; neuroprotective; vasotropic; cytostatic;
KW   immunosuppressive; inhibitor of apoptosis protein; IAP;
KW   caspase-9 N-terminal linker; procaspase-9; cysteine protease; caspase-3;
```

KW Bir3 domain; apoptosis; AIDS; neurodegenerative disease;
KW ischaemic injury; cancer; autoimmune disease; Fruitfly; Grim.
XX
OS Drosophila melanogaster.
XX
PN US2002160975-A1.
XX
PD 31-OCT-2002.
XX
PF 06-FEB-2002; 2002US-00068569.
XX
PR 08-FEB-2001; 2001US-0267966P.
PR 24-AUG-2001; 2001US-00939293.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Alnemri ES;
XX
DR WPI; 2003-219992/21.
XX
PT New nucleic acid molecules encoding a peptide or polypeptide that binds
PT to a portion of an inhibitor of apoptosis protein, useful for inducing
PT apoptosis and identifying inhibitors or enhancers of apoptosis for
PT treating AIDS, or cancer.
XX
PS Claim 27; Fig 9; 52pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule comprising a
CC polynucleotide that encodes a polypeptide or peptide, or its variants
CC that specifically binds to at least a portion of an inhibitor of
CC apoptosis protein (IAP). Also included are a peptide or a polypeptide
CC (comprising at least an N terminus sequence of caspase-9 N-terminal
CC linker sequence, a first portion of a procaspase-9 that specifically
CC binds at least a portion of an IAP and a second portion of a procaspase-9
CC containing a mutated active site, where the peptide or polypeptide
CC specifically binds at least a portion of an IAP and lacks cysteine
CC protease activity, and at least a portion of caspase-3, where the peptide
CC or polypeptide exhibits caspase-3 enzymatic activity that is inhibited by
CC an IAP or an IAP Bir3 domain) or at least a portion of a mutated
CC procaspase-9, which fails to undergo normal processing and possesses wild
CC type caspase-9 enzymatic activity, a nucleic acid molecule comprising a
CC polynucleotide sequence that encodes the caspase-9 N-terminal linker), an
CC expression vector comprising any of the nucleic acids, a host cell
CC containing the expression vector, an antibody that specifically binds to
CC the peptide or polypeptide, an antibody that specifically binds to an
CC epitope located on the N-terminus of a caspase-9-p12, inducing apoptosis
CC in a cell or stimulating apoptosis in a neoplastic or tumour cell,
CC identifying an inhibitor or enhancer of caspase-mediated apoptosis,
CC identifying a compound that inhibits the peptide or polypeptide,
CC producing a compound for inhibiting or enhancing apoptosis in a cell, and
CC a process for the manufacture of a compound for inhibiting or enhancing
CC apoptosis in a cell. The nucleic acid molecules and peptides or
CC polypeptides are useful for inducing apoptosis and identifying inhibitors
CC or enhancers of apoptosis for treating AIDS, neurodegenerative diseases,
CC ischaemic injury, cancer, autoimmune diseases. The present sequence
CC represents the N-terminus of an IAP protein containing a Bir3 domain.
XX
SQ Sequence 15 AA;

```
  Query Match             40.8%;  Score 31;  DB 6;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 2.3e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 LVHFVRDFAQLL 15
              : :|: | ||||
Db          2 IAYFIPDQAQLL 13
```

RESULT 7
AAY06719
ID AAY06719 standard; peptide; 14 AA.
XX
AC AAY06719;
XX
DT 18-JUN-1999 (first entry)
XX
DE Peptide fragment from bovine prion protein (residues 146-159).
XX
KW Prion; PrP-Sc protein; transmissible spongiform encephalopathy; TSE;
KW ruminant; immunoassay; third eyelid-associated lymphoid tissue; epitope;
KW neurodegenarative disorder; scrapie; bovine prion protein.
XX
OS Synthetic.
OS Bos sp.
XX
PN WO9919360-A1.
XX
PD 22-APR-1999.
XX
PF 14-OCT-1998; 98WO-US021704.
XX
PR 14-OCT-1997; 97US-00950271.
XX
PA (USDA) US SEC OF AGRIC.
PA (UNIW) UNIV WASHINGTON STATE RES FOUND.
XX
PI O'rourke KI, Knowles DP, Baszler TV, Parish SM;
XX
DR WPI; 1999-277597/23.
XX
PT Assay for detecting prion (PrP-Sc) protein - as indicator for
PT transmissible spongiform encephalopathy.
XX
PS Example 2; Page 15; 33pp; English.
XX
CC The invention relates to immunoassay for detecting the presence of PrP-Sc
CC (prion) protein as an indication of transmissible spongiform
CC encephalopathy (TSE) in a ruminant. The assay comprises (a) obtaining
CC third eyelid-associated lymphoid tissue from a ruminant to be tested; (b)
CC treating the tissue to unmask an epitope to PrP-Sc protein and eliminate
CC availability of a corresponding epitope of PRP-C which is expressed in
CC tissues from normal animals; (c) contacting said treated tissue with an
CC antibody which specifically binds said PrP-Sc epitope under conditions
CC such that said antibody binds PrP-Sc protein if said protein is present
CC in the tissue, and (d) detecting the presence of said bound antibody. The
CC methods are used for the detection of transmissible spongioform
CC encephalopathies which are a heterogenous group of fatal
CC neurodegenarative disorders occurring in humans, ruminants, herbivores,
CC mink and cats, sheep scrapie being a representative of this group. The
CC methods are non-invasive and pre-clinical. The present sequence
CC represents a peptide fragment from the bovine prion protein
XX
SQ Sequence 14 AA;

```
  Query Match             39.5%;  Score 30;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RQLVHFVRDF 11
              | |:||  |:
Db          2 RPLIHFGSDY 11
```

RESULT 8
ADC21478
ID ADC21478 standard; peptide; 9 AA.
XX
AC ADC21478;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human PRDI-BF1 peptide fragment SEQ ID 15.
XX
KW tumor; antigen; CD8+ cytotoxic T lymphocyte; CTL; CTL-induced lysis;
KW multiple myeloma cell; human; PRDI-BF1;
KW positive regulatory domain I-binding factor-1; MHC;

KW major histocompatibility complex Class I; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003029282-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-EP010701.
XX
PR 29-SEP-2001; 2001DE-01048236.
XX
PA (IMMU-) IMMUGENICS AG.
XX
PI Theobald M, Lotz C;
XX
DR WPI; 2003-354724/33.
XX
PT New tumor-associated oligopeptide, useful particularly for treating
PT multiple myeloma, is recognized by CD8 cytotoxic T cells, also
PT derivatives and related nucleic acid.
XX
PS Claim 1; SEQ ID NO 15; 64pp; German.
XX
CC This invention describes a novel tumor-associated oligopeptide that is
CC recognized as an antigen by CD8+ cytotoxic T lymphocytes (CTL) and causes
CC CTL-induced lysis and/or apoptosis of tumor cells, especially multiple
CC myeloma cells. The oligopeptide is derived from human PRDI-BF1 (positive
CC regulatory domain I-binding factor-1) which is able to induce an MHC
CC (major histocompatibility complex) Class I allele variant A2-restricted
CC immune response of CD8+ CTL against tumor cells. The products of the
CC invention have cytostatic activity and can be used in a vaccine. The
CC peptide of the invention, also related retro-inverse and pseudopeptides,
CC fusion proteins (FP), polynucleotides, vectors, host cells and antibodies
CC and T cell receptors specific for PRDI-BFI peptides are useful for
CC treating diseases associated with PRDI-BF1, particularly tumors. The
CC products of the invention are also useful as diagnostic, therapeutic and
CC prophylactic agents for detecting, modifying, generating, expanding
CC and/or regulating activation and functional status of T cells, and for
CC preparation of poly- or mono-clonal or recombinant A2-restricted T cell
CC receptors and their functional equivalents.
XX
SQ Sequence 9 AA;

```
  Query Match             38.2%;  Score 29;  DB 7;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 LVHFVRDFA 12
              || : ||||
Db          1 LVWYCRDFA 9
```

RESULT 9
ADC21483
ID ADC21483 standard; peptide; 10 AA.
XX
AC ADC21483;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human PRDI-BF1 peptide fragment SEQ ID 20.
XX
KW tumor; antigen; CD8+ cytotoxic T lymphocyte; CTL; CTL-induced lysis;
KW multiple myeloma cell; human; PRDI-BF1;
KW positive regulatory domain I-binding factor-1; MHC;
KW major histocompatibility complex Class I; cytostatic; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003029282-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-EP010701.
XX
PR 29-SEP-2001; 2001DE-01048236.
XX
PA (IMMU-) IMMUGENICS AG.
XX
PI Theobald M, Lotz C;
XX
DR WPI; 2003-354724/33.
XX
PT New tumor-associated oligopeptide, useful particularly for treating
PT multiple myeloma, is recognized by CD8 cytotoxic T cells, also
PT derivatives and related nucleic acid.
XX
PS Claim 1; SEQ ID NO 20; 64pp; German.
XX
CC This invention describes a novel tumor-associated oligopeptide that is
CC recognized as an antigen by CD8+ cytotoxic T lymphocytes (CTL) and causes
CC CTL-induced lysis and/or apoptosis of tumor cells, especially multiple
CC myeloma cells. The oligopeptide is derived from human PRDI-BF1 (positive
CC regulatory domain I-binding factor-1) which is able to induce an MHC
CC (major histocompatibility complex) Class I allele variant A2-restricted
CC immune response of CD8+ CTL against tumor cells. The products of the
CC invention have cytostatic activity and can be used in a vaccine. The
CC peptide of the invention, also related retro-inverse and pseudopeptides,
CC fusion proteins (FP), polynucleotides, vectors, host cells and antibodies
CC and T cell receptors specific for PRDI-BFI peptides are useful for
CC treating diseases associated with PRDI-BF1, particularly tumors. The
CC products of the invention are also useful as diagnostic, therapeutic and
CC prophylactic agents for detecting, modifying, generating, expanding
CC and/or regulating activation and functional status of T cells, and for
CC preparation of poly- or mono-clonal or recombinant A2-restricted T cell
CC receptors and their functional equivalents.
XX
SQ Sequence 10 AA;

```
  Query Match             38.2%;  Score 29;  DB 7;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 3.4e+02;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 LVHFVRDFA 12
              || : ||||
Db          2 LVWYCRDFA 10
```

RESULT 10
AAG84689
ID AAG84689 standard; peptide; 9 AA.
XX
AC AAG84689;
XX
DT 10-SEP-2001 (first entry)
XX
DE MAGE3 crossbinding data A2 supermotif peptide #13.
XX
KW Human; human leukocyte antigen; HLA epitope; cytotoxic T lymphocyte; CTL;
KW MAGE2; MAGE3; melanoma antigen gene; immune response; vaccine; cancer;
KW cytostatic; immunostimulant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200142267-A1.
XX
PD 14-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US033545.
XX
PR 10-DEC-1999; 99US-00458298.
XX
PA (EPIM-) EPIMMUNE INC.
XX

PI Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;
PI Keogh E;
XX
DR WPI; 2001-375002/39.
XX
PT An isolated prepared MAGE2/3 epitope (I) for use in pharmaceuticals for
PT the treatment and prevention of cancer.
XX
PS Example 2; Page 152; 171pp; English.
XX
CC The present invention describes MAGE2/3 epitopes (I). Also described are:
CC (1) a clonal cytotoxic T lymphocyte (CTL) that is cultured in vivo and
CC binds to a complex of (I); (2) a peptide (II) comprising (I) and a second
CC epitope and has less than 50 contiguous amino acids; (3) a vaccine
CC composition comprising (II), a unit dose of a peptide with at least 50
CC contiguous amino acids with 100% identity to the native peptide sequence
CC of MAGE2/3, and a pharmaceutical excipient; (4) an isolated nucleic acid
CC encoding (I); and (5) an isolated nucleic acid encoding (II). (I) has
CC cytostatic activity, and can be used in vaccines and as an
CC immunostimulant. A vaccine of (3) is useful for the treatment and
CC prevention of cancer. (I) is useful for monitoring or evaluating an
CC immune response by incubating a T-lymphocyte sample from a patient with
CC (I) that binds to an human leukocyte antigen (HLA) allele present in the
CC patient and detecting the presence of the T-lymphocyte that binds to the
CC peptide. The vaccine allows the opportunity to combine epitopes derived
CC from multiple tumour-associated molecules reducing the likelihood of
CC tumour escape due to antigen loss. AAG84515 to AAG84909 and AAB99725
CC represent amino acid sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 9 AA;

Query Match 36.8%; Score 28; DB 4; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QLVHFV 8
:|||||
Db 4 ELVHFV 9

RESULT 11
AAG84688
ID AAG84688 standard; peptide; 9 AA.
XX
AC AAG84688;
XX
DT 10-SEP-2001 (first entry)
XX
DE MAGE3 crossbinding data A2 supermotif peptide #12.
XX
KW Human; human leukocyte antigen; HLA epitope; cytotoxic T lymphocyte; CTL;
KW MAGE2; MAGE3; melanoma antigen gene; immune response; vaccine; cancer;
KW cytostatic; immunostimulant.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200142267-A1.
XX
PD 14-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US033545.
XX
PR 10-DEC-1999; 99US-00458298.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Fikes J, Sette A, Sidney J, Southwood S, Chesnut R, Celis E;
PI Keogh E;
XX
DR WPI; 2001-375002/39.
XX
PT An isolated prepared MAGE2/3 epitope (I) for use in pharmaceuticals for
PT the treatment and prevention of cancer.
XX
PS Example 2; Page 152; 171pp; English.
XX
CC The present invention describes MAGE2/3 epitopes (I). Also described are:
CC (1) a clonal cytotoxic T lymphocyte (CTL) that is cultured in vivo and
CC binds to a complex of (I); (2) a peptide (II) comprising (I) and a second
CC epitope and has less than 50 contiguous amino acids; (3) a vaccine
CC composition comprising (II), a unit dose of a peptide with at least 50
CC contiguous amino acids with 100% identity to the native peptide sequence
CC of MAGE2/3, and a pharmaceutical excipient; (4) an isolated nucleic acid
CC encoding (I); and (5) an isolated nucleic acid encoding (II). (I) has
CC cytostatic activity, and can be used in vaccines and as an
CC immunostimulant. A vaccine of (3) is useful for the treatment and
CC prevention of cancer. (I) is useful for monitoring or evaluating an
CC immune response by incubating a T-lymphocyte sample from a patient with
CC (I) that binds to an human leukocyte antigen (HLA) allele present in the
CC patient and detecting the presence of the T-lymphocyte that binds to the
CC peptide. The vaccine allows the opportunity to combine epitopes derived
CC from multiple tumour-associated molecules reducing the likelihood of
CC tumour escape due to antigen loss. AAG84515 to AAG84909 and AAB99725
CC represent amino acid sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 9 AA;

Query Match 36.8%; Score 28; DB 4; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QLVHFV 8
:|||||
Db 4 ELVHFV 9

RESULT 12
AAU27075
ID AAU27075 standard; peptide; 9 AA.
XX
AC AAU27075;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human Leukocyte Antigen (HLA) HLA-A2.1 immunogenic binding peptide #359.
XX
KW Immunogenic peptide; human leukocyte antigen; HLA-A2.1 binding motif;
KW immunostimulant; cytostatic; antiviral; glycoprotein; cytotoxic T cell;
KW viral disease; prostate cancer; hepatitis B; hepatitis C; lymphoma; AIDS;
KW renal carcinoma; cervical carcinoma; condyloma acuminatum.
XX
OS Homo sapiens.
XX
PN WO200162776-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2000; 2000WO-US004655.
XX
PR 23-FEB-2000; 2000WO-US004655.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Kast WM, Southwood S;
XX
DR WPI; 2001-582039/65.
XX
PT Composition for treating viral diseases and cancer comprises an
PT immunogenic peptide having an HLA-A2.1 binding motif.
XX
PS Claim 1; Page 71; 85pp; English.
XX
CC Sequences AAU26558-AAU27161 represent immunogenic peptides containing a

CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
CC the invention are capable of specifically binding glycoproteins encoded
CC by HLA alleles and inducing a cytotoxic T cell response against an
CC antigen in a patient expressing HLA-A2.1. This method is useful for the
CC treatment, prevention and diagnosis of pathological states such as viral
CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis
CC C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma
CC acuminatum. The peptides are used for treatment of chronic infection and
CC for stimulating the immune system to eliminate virus-infected cells
XX
SQ Sequence 9 AA;

Query Match 36.8%; Score 28; DB 4; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QLVHFV 8
:|||||
Db 4 ELVHFV 9

RESULT 13
AAU27076
ID AAU27076 standard; peptide; 9 AA.
XX
AC AAU27076;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human Leukocyte Antigen (HLA) HLA-A2.1 immunogenic binding peptide #360.
XX
KW Immunogenic peptide; human leukocyte antigen; HLA-A2.1 binding motif;
KW immunostimulant; cytostatic; antiviral; glycoprotein; cytotoxic T cell;
KW viral disease; prostate cancer; hepatitis B; hepatitis C; lymphoma; AIDS;
KW renal carcinoma; cervical carcinoma; condyloma acuminatum.
XX
OS Homo sapiens.
XX
PN WO200162776-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2000; 2000WO-US004655.
XX
PR 23-FEB-2000; 2000WO-US004655.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Kast WM, Southwood S;
XX
DR WPI; 2001-582039/65.
XX
PT Composition for treating viral diseases and cancer comprises an
PT immunogenic peptide having an HLA-A2.1 binding motif.
XX
PS Claim 1; Page 71; 85pp; English.
XX
CC Sequences AAU26558-AAU27161 represent immunogenic peptides containing a
CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
CC the invention are capable of specifically binding glycoproteins encoded
CC by HLA alleles and inducing a cytotoxic T cell response against an
CC antigen in a patient expressing HLA-A2.1. This method is useful for the
CC treatment, prevention and diagnosis of pathological states such as viral
CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis
CC C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma
CC acuminatum. The peptides are used for treatment of chronic infection and
CC for stimulating the immune system to eliminate virus-infected cells
XX
SQ Sequence 9 AA;

Query Match 36.8%; Score 28; DB 4; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QLVHFV 8
:|||||
Db 4 ELVHFV 9

RESULT 14
AAU26742
ID AAU26742 standard; peptide; 9 AA.
XX
AC AAU26742;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human Leukocyte Antigen (HLA) HLA-A2.1 immunogenic binding peptide #185.
XX
KW Immunogenic peptide; human leukocyte antigen; HLA-A2.1 binding motif;
KW immunostimulant; cytostatic; antiviral; glycoprotein; cytotoxic T cell;
KW viral disease; prostate cancer; hepatitis B; hepatitis C; lymphoma; AIDS;
KW renal carcinoma; cervical carcinoma; condyloma acuminatum.
XX
OS Homo sapiens.
XX
PN WO200162776-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2000; 2000WO-US004655.
XX
PR 23-FEB-2000; 2000WO-US004655.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Kast WM, Southwood S;
XX
DR WPI; 2001-582039/65.
XX
PT Composition for treating viral diseases and cancer comprises an
PT immunogenic peptide having an HLA-A2.1 binding motif.
XX
PS Example 1; Page 33; 85pp; English.
XX
CC Sequences AAU26558-AAU27161 represent immunogenic peptides containing a
CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
CC the invention are capable of specifically binding glycoproteins encoded
CC by HLA alleles and inducing a cytotoxic T cell response against an
CC antigen in a patient expressing HLA-A2.1. This method is useful for the
CC treatment, prevention and diagnosis of pathological states such as viral
CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis
CC C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma
CC acuminatum. The peptides are used for treatment of chronic infection and
CC for stimulating the immune system to eliminate virus-infected cells
XX
SQ Sequence 9 AA;

Query Match 36.8%; Score 28; DB 4; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 QLVHFV 8
:|||||
Db 4 ELVHFV 9

RESULT 15
AAU26743
ID AAU26743 standard; peptide; 9 AA.
XX
AC AAU26743;
XX
DT 18-DEC-2001 (first entry)
XX
DE Human Leukocyte Antigen (HLA) HLA-A2.1 immunogenic binding peptide #186.

XX
KW Immunogenic peptide; human leukocyte antigen; HLA-A2.1 binding motif;
KW immunostimulant; cytostatic; antiviral; glycoprotein; cytotoxic T cell;
KW viral disease; prostate cancer; hepatitis B; hepatitis C; lymphoma; AIDS;
KW renal carcinoma; cervical carcinoma; condyloma acuminatum.
XX
OS Homo sapiens.
XX
PN WO200162776-A1.
XX
PD 30-AUG-2001.
XX
PF 23-FEB-2000; 2000WO-US004655.
XX
PR 23-FEB-2000; 2000WO-US004655.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Kast WM, Southwood S;
XX
DR WPI; 2001-582039/65.
XX
PT Composition for treating viral diseases and cancer comprises an
PT immunogenic peptide having an HLA-A2.1 binding motif.
XX
PS Example 1; Page 33; 85pp; English.
XX
CC Sequences AAU26558-AAU27161 represent immunogenic peptides containing a
CC human leukocyte antigen A2.1 (HLA-A2.1) binding motif. The peptides of
CC the invention are capable of specifically binding glycoproteins encoded
CC by HLA alleles and inducing a cytotoxic T cell response against an
CC antigen in a patient expressing HLA-A2.1. This method is useful for the
CC treatment, prevention and diagnosis of pathological states such as viral
CC diseases and cancers, including prostate cancer, hepatitis B, hepatitis
CC C, AIDS, renal carcinoma, cervical carcinoma, lymphoma, and condyloma
CC acuminatum. The peptides are used for treatment of chronic infection and
CC for stimulating the immune system to eliminate virus-infected cells
XX
SQ Sequence 9 AA;

Query Match 36.8%; Score 28; DB 4; Length 9;
Best Local Similarity 83.3%; Pred. No. 1.8e+06;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 QLVHFV 8
              :|||||
Db          4 ELVHFV 9
```

Search completed: February 22, 2005, 09:24:44
Job time : 67.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-9
Perfect score: 77
Sequence: 1 DFEVVTFLKDVLPEF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	23	29.9	15	2	G49732	NADH2 dehydrogenas
2	22	28.6	14	2	E33098	214K exoantigen (v
3	22	28.6	15	2	A36315	recycling receptor
4	22	28.6	15	2	PN0662	dystrophin-associa
5	22	28.6	15	2	PA0061	protein QF200039 -
6	21	27.3	8	2	S66296	Na+-transporting A
7	21	27.3	15	2	PS0185	27K protein A 3.4/
8	20	26.0	9	2	S78426	52.5K protein - sp
9	20	26.0	11	2	S41747	chaperonin 10 homo
10	20	26.0	12	2	PT0228	Ig heavy chain CDR
11	20	26.0	14	2	I54945	gene C protein - E
12	20	26.0	14	2	S14336	mastoparan B - hor
13	20	26.0	15	2	B61457	alpha-glucosidase
14	19	24.7	7	2	B39127	phosphotransferase
15	19	24.7	14	2	PS0249	porin - rice (stra
16	19	24.7	15	2	T09463	ribosomal protein
17	19	24.7	15	2	S62675	collagen type I -
18	18	23.4	9	2	B30572	T-cell receptor be
19	18	23.4	9	2	S77984	cytochrome-c oxida
20	18	23.4	10	2	C30572	T-cell receptor be
21	18	23.4	12	2	PS0213	28K protein 4412 -
22	18	23.4	13	2	PN0122	O1L protein - vacc
23	18	23.4	14	2	S29789	hypothetical prote
24	18	23.4	14	2	A60737	pollen allergen Lo
25	18	23.4	14	2	PC4382	dehydrin 4.5K poly
26	18	23.4	14	2	A44920	2-halobenzoate 1,2
27	18	23.4	14	2	PC4376	telomeric and tetr
28	18	23.4	15	2	PH1807	T cell receptor al
29	18	23.4	15	2	A26228	spot 42 protein -
30	17.5	22.7	15	2	PQ0195	Sf11-glycoprotein
31	17	22.1	10	2	S48182	bacterioferritin -
32	17	22.1	10	2	A44646	neurotoxin-associa
33	17	22.1	11	2	PT0229	Ig heavy chain CDR
34	17	22.1	11	2	PT0249	Ig heavy chain CRD
35	17	22.1	11	2	PD0441	translation elonga
36	17	22.1	12	2	PA0047	protein QA100045 -
37	17	22.1	12	2	I58273	thyroglobulin - ra
38	17	22.1	13	2	PH1599	Ig H chain V-D-J r
39	17	22.1	14	2	S23369	T-cell receptor al
40	17	22.1	14	2	A61032	troponin T, cardia
41	17	22.1	14	2	A41589	25K elastin-bindin
42	17	22.1	14	2	PC7075	guanylate cyclase
43	17	22.1	15	2	S29485	GTP-binding protei
44	17	22.1	15	2	PH1612	Ig H chain V-D-J r
45	17	22.1	15	2	PA0093	enniatin synthetas

ALIGNMENTS

RESULT 1
G49732
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) 20K chain - potato mitochondrion (fragment)
N;Alternate names: complex 1 hydrogenase 20K chain; NADH-ubiquinone oxidoreductase 20K ch
C;Species: mitochondrion Solanum tuberosum (potato)
C;Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 09-Jul-2004
C;Accession: G49732
R;Herz, U.; Schroeder, W.; Liddell, A.; Leaver, C.J.; Brennicke, A.; Grohmann, L.
J. Biol. Chem. 269, 2263-2269, 1994
A;Title: Purification of the NADH:ubiquinone oxidoreductase (complex I) of the respirator
A;Reference number: A49732; MUID:94124587; PMID:8294484
A;Accession: G49732
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <HER>
A;Cross-references: UNIPROT:P80263
C;Genetics:
A;Genome: mitochondrion
C;Keywords: electron transfer; hydrogen ion transport; mitochondrial inner membrane; mito

Query Match 29.9%; Score 23; DB 2; Length 15;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          6 TFLKDV 11
              | ||||
Db          8 TXLKDV 13
```

RESULT 2
E33098
214K exoantigen (version 2) - malaria parasite (Plasmodium falciparum) (fragments)
C;Species: Plasmodium falciparum
C;Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000
C;Accession: E33098
R;Nichols, J.H.; Hager, L.P.
submitted to the Protein Sequence Database, May 1990
A;Reference number: A33098
A;Accession: E33098
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <NIC>

Query Match 28.6%; Score 22; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 1.9e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy          3 EVVTFL 8
              :|||:|
Db          5 DVVTYL 10
```

RESULT 3
A36315
recycling receptor p180 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 25-Jan-1991 #sequence_revision 25-Jan-1991 #text_change 30-Sep-1993
C;Accession: A36315
R;Isacke, C.M.; van der Geer, P.; Hunter, T.; Trowbridge, I.S.
Mol. Cell. Biol. 10, 2606-2618, 1990
A;Title: p180, a novel recycling transmembrane glycoprotein with restricted cell type ex
A;Reference number: A36315; MUID:90258846; PMID:2188094
A;Accession: A36315
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <ISA>

Query Match 28.6%; Score 22; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 10 DVLPE 14
| |||
Db 5 DALPE 9

RESULT 4
PN0662
dystrophin-associated glycoprotein A3a-I - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 07-May-1999
C;Accession: PN0662
R;Yoshida, M.; Mizuno, Y.; Nonaka, I.; Ozawa, E.
J. Biochem. 114, 634-639, 1993
A;Title: A dystrophin-associated glycoprotein, A3a (one of 43DAG doublets), is retained
A;Reference number: PN0662; MUID:94156881; PMID:8113213
A;Accession: PN0662
A;Molecule type: protein
A;Residues: 1-15 <YOS>
C;Comment: This protein is retained in Duchenne type muscular dystrophy muscle.
C;Keywords: glycoprotein; skeletal muscle

Query Match 28.6%; Score 22; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 2e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 6 TFLKDVLP 13
||:| :|
Db 4 TFIKKGVP 11

RESULT 5
PA0061
protein QF200039 - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0061
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0061
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Y2

Query Match 28.6%; Score 22; DB 2; Length 15;
Best Local Similarity 44.4%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 5 VTFLKDVLP 13
| :|| |
Db 7 VAIVKDTAP 15

RESULT 6
S66296
Na+-transporting ATP synthase (EC 3.6.1.-) chain c - Acetobacterium woodii (fragment)
N;Alternate names: ATPase chain c
C;Species: Acetobacterium woodii
C;Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 07-May-1999
C;Accession: S66296
R;Reidlinger, J.; Mueller, V.
Eur. J. Biochem. 223, 275-283, 1994
A;Title: Purification of ATP synthase from Acetobacterium woodii and identification as a
A;Reference number: S45648; MUID:94307271; PMID:8033902
A;Accession: S66296
A;Molecule type: protein
A;Residues: 1-8 <REI>
A;Experimental source: DSM 1030
C;Keywords: hydrolase

Query Match 27.3%; Score 21; DB 2; Length 8;
Best Local Similarity 42.9%; Pred. No. 2.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 EVVTFLK 9
|:: |:|
Db 2 EILDFIK 8

RESULT 7
PS0185
27K protein A 3.4/5 - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 24-Feb-1995
C;Accession: PS0185
R;Kamo, M.; Tsugita, A.
submitted to JIPID, June 1991
A;Reference number: PS0184
A;Accession: PS0185
A;Molecule type: protein
A;Residues: 1-15 <KAM>

Query Match 27.3%; Score 21; DB 2; Length 15;
Best Local Similarity 37.5%; Pred. No. 2.9e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 7 FLKDVLPE 14
:: || |:
Db 4 YIVDVAPD 11

RESULT 8
S78426
52.5K protein - spiny lobster (fragment)
C;Species: Panulirus argus (spiny lobster)
C;Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 19-May-2000
C;Accession: S78426
R;James, M.O.; Boyle, S.M.; Trapido-Rosenthal, H.G.; Smith, W.C.; Greenberg, R.M.; Shive
Arch. Biochem. Biophys. 329, 31-38, 1996
A;Title: cDNA and protein sequence of a major form of P450, CYP2L, in the hepatopancreas
A;Reference number: S68856; MUID:96201120; PMID:8619632
A;Accession: S78426
A;Molecule type: protein
A;Residues: 1-9 <JAM>
A;Experimental source: hepatopancreas microsomes

Query Match 26.0%; Score 20; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 TFLKDVL 12
|::| ||
Db 1 TWIKKVL 7

RESULT 9

S41747
chaperonin 10 homolog - potato (fragment)
C;Species: Solanum tuberosum (potato)
C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 09-Jul-2004
C;Accession: S41747
R;Burt, W.J.E.; Leaver, C.J.
FEBS Lett. 339, 139-141, 1994
A;Title: Identification of a chaperonin-10 homologue in plant mitochondria.
A;Reference number: S41747; MUID:94148071; PMID:7906228
A;Accession: S41747
A;Molecule type: protein
A;Residues: 1-11 <BUR>
A;Cross-references: UNIPROT:Q7M1H1
A;Experimental source: mitochondrion
C;Keywords: mitochondrion; molecular chaperone

Query Match 26.0%; Score 20; DB 2; Length 11;
Best Local Similarity 60.0%; Pred. No. 3.1e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 11 VLPEF 15
:|||:
Db 2 LLPEY 6

RESULT 10
PT0228
Ig heavy chain CDR3 region (clone 1-112) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0228
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0228
A;Molecule type: DNA
A;Residues: 1-12 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 26.0%; Score 20; DB 2; Length 12;
Best Local Similarity 37.5%; Pred. No. 3.4e+03;
Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 7 FLKDVLPE 14
:::| ||
Db 3 YVRDSSPE 10

RESULT 11
I54945
gene C protein - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 07-Jun-1996 #sequence_revision 07-Jun-1996 #text_change 09-Jul-2004
C;Accession: I54945
R;Tao, T.; Bourne, J.C.; Blumenthal, R.M.
J. Bacteriol. 173, 1367-1375, 1991
A;Title: A family of regulatory genes associated with type II restriction-modification s
A;Reference number: I54945; MUID:91139577; PMID:1995588
A;Accession: I54945
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-14 <RES>
A;Cross-references: UNIPROT:Q47599; GB:M63619; NID:g147664; PIDN:AAA24555.1; PID:g147665

Query Match 26.0%; Score 20; DB 2; Length 14;
Best Local Similarity 57.1%; Pred. No. 4e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 8 LKDVLPE 14
||:|: |
Db 1 LKEVIME 7

RESULT 12
S14336
mastoparan B - hornet (Vespa basalis)
C;Species: Vespa basalis
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
C;Accession: S14336
R;Ho, C.L.; Hwang, L.L.
Biochem. J. 274, 453-456, 1991
A;Title: Structure and biological activities of a new mastoparan isolated from the venom
A;Reference number: S14336; MUID:91174755; PMID:2006909
A;Accession: S14336
A;Molecule type: protein
A;Residues: 1-14 <HOC>
A;Cross-references: UNIPROT:P21654
A;Experimental source: venom
C;Function:
A;Description: posesses a potent hemolytic activity which acts in synergy with the lethal
C;Keywords: amidated carboxyl end; mast cell; venom
F;14/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 26.0%; Score 20; DB 2; Length 14;
Best Local Similarity 44.4%; Pred. No. 4e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 VVTFLKDVL 12
:|:: | ||
Db 6 IVSWAKKVL 14

RESULT 13
B61457
alpha-glucosidase (EC 3.2.1.20) - Tetrahymena pyriformis (strain W) (fragment)
C;Species: Tetrahymena pyriformis
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 07-Dec-1999
C;Accession: B61457
R;Banno, Y.; Sasaki, N.; Yoshino, T.; Mochizuki, J.I.; Hirata, H.; Nozawa, Y.
J. Protozool. 36, 562-567, 1989
A;Title: A thermostable acid alpha-glucosidase from Tetrahymena thermophila: purificatior
A;Reference number: A61457; MUID:90095988; PMID:2689637
A;Accession: B61457
A;Molecule type: protein
A;Residues: 1-15 <BAN>
C;Genetics:
A;Genetic code: SGC5
C;Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; lysosome; monome

Query Match 26.0%; Score 20; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 4.3e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 VLPEF 15
||| |
Db 1 VLPPF 5

RESULT 14
B39127
phosphotransferase system enzyme II (EC 2.7.1.69) - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 08-Oct-1999
C;Accession: B39127
R;Hardesty, C.; Ferran, C.; DiRienzo, J.M.
J. Bacteriol. 173, 449-456, 1991
A;Title: Plasmid-mediated sucrose metabolism in Escherichia coli: characterization of scr
rin.
A;Reference number: A39127; MUID:91100329; PMID:1846143
A;Accession: B39127
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-7 <HAR>
A;Cross-references: GB:M38416; NID:g155142; PIDN:AAA98418.1; PID:g155144

C;Keywords: phosphotransferase

Query Match 24.7%; Score 19; DB 2; Length 7;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 DFEVVT 6
              ||| ::
Db          2 DFEQIS 7
```

RESULT 15
PS0249
porin - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: PS0249
R;Tsugita, A.
submitted to JIPID, April 1993
A;Reference number: PS0206
A;Accession: PS0249
A;Molecule type: protein
A;Residues: 1-14 <TSU>
A;Cross-references: UNIPROT:Q7M1U8
A;Experimental source: callus

Query Match 24.7%; Score 19; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 5.8e+03;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          5 VTFLKD 10
              |||  |
Db          2 VTFTDD 7
```

Search completed: February 22, 2005, 09:46:27
Job time : 11.1333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-9
Perfect score: 77
Sequence: 1 DFEVVTFLKDVLPEF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	33.8	13	2	Q6TKD3	Q6tkd3 praecitrull
2	26	33.8	13	2	Q6TKD4	Q6tkd4 sechium edu
3	26	33.8	13	2	Q6TKD5	Q6tkd5 sicyos angu
4	26	33.8	13	2	Q6TKD6	Q6tkd6 trichosanth
5	26	33.8	13	2	Q6TKD7	Q6tkd7 luffa grave
6	26	33.8	13	2	Q6TKD8	Q6tkd8 luffa echin
7	26	33.8	13	2	Q6TKD9	Q6tkd9 cucurbita p
8	26	33.8	13	2	Q6TKE0	Q6tke0 benincasa h
9	26	33.8	13	2	Q6TKE1	Q6tke1 marah orega
10	26	33.8	13	2	Q6TKE2	Q6tke2 cyclanthera
11	26	33.8	13	2	Q6TKE3	Q6tke3 echinocysti
12	26	33.8	13	2	Q6TKE4	Q6tke4 luffa quinq
13	26	33.8	13	2	Q6TKE5	Q6tke5 lagenaria l
14	26	33.8	13	2	Q6TKE6	Q6tke6 citrullus l
15	26	33.8	13	2	Q6TKE7	Q6tke7 citrullus c
16	26	33.8	13	2	Q6TKE8	Q6tke8 acanthosicy
17	26	33.8	13	2	Q6TKE9	Q6tke9 bryonia dio
18	26	33.8	13	2	Q6TKF0	Q6tkf0 coccinia pa
19	26	33.8	13	2	Q6TKF1	Q6tkf1 diplocyclos
20	26	33.8	13	2	Q6TKF2	Q6tkf2 ecballium e
21	26	33.8	15	2	Q9R565	Q9r565 streptomyce
22	24	31.2	10	2	P82222	P82222 bombyx mori
23	24	31.2	12	2	P82246	P82246 spinacia ol
24	24	31.2	12	2	P82247	P82247 spinacia ol
25	23	29.9	10	1	AKHX_LOCMI	P81626 locusta mig
26	23	29.9	12	2	Q7RH69	Q7rh69 plasmodium
27	23	29.9	14	2	O70599	O70599 rattus norv
28	23	29.9	15	1	NUO3_SOLTU	P80263 solanum tub
29	23	29.9	15	2	Q9UCJ8	Q9ucj8 homo sapien
30	23	29.9	15	2	Q9R599	Q9r599 micrococcus
31	22	28.6	14	2	Q7Z7E2	Q7z7e2 homo sapien
32	22	28.6	15	2	Q7M4Y2	Q7m4y2 fusarium sp
33	22	28.6	15	2	Q7XQX8	Q7xqx8 oryza sativ
34	21	27.3	13	2	Q8I2E2	Q8i2e2 plasmodium
35	21	27.3	14	1	RAN9_RANCA	P82824 rana catesb
36	20	26.0	9	2	Q7RSP1	Q7rsp1 plasmodium
37	20	26.0	10	2	Q6JDL6	Q6jdl6 canis famil
38	20	26.0	11	2	Q7M1H1	Q7m1h1 solanum tub
39	20	26.0	12	1	XYLA_STRVN	P14405 streptomyce
40	20	26.0	13	2	Q16007	Q16007 homo sapien
41	20	26.0	13	2	Q83171	Q83171 cauliflower
42	20	26.0	13	2	Q98YM1	Q98ym1 human immun
43	20	26.0	14	1	MAST_VESBA	P21654 vespa basal
44	20	26.0	14	2	Q47599	Q47599 escherichia
45	20	26.0	15	1	UC16_MAIZE	P80622 zea mays (m

ALIGNMENTS

RESULT 1
Q6TKD3
ID Q6TKD3 PRELIMINARY; PRT; 13 AA.
AC Q6TKD3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Photosystem II protein (Fragment).
GN Name=psbC;
OS Praecitrullus fistulosus.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Praecitrullus.
OX NCBI_TaxID=252558;
RN [1]
RP SEQUENCE FROM N.A.
RA Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY396191; AAR07576.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
KW Chloroplast.
FT NON_TER 1 1
SQ SEQUENCE 13 AA; 1579 MW; CF0270FBB503C732 CRC64;

Query Match 33.8%; Score 26; DB 2; Length 13;
Best Local Similarity 62.5%; Pred. No. 2e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9
```

RESULT 2
Q6TKD4
ID Q6TKD4 PRELIMINARY; PRT; 13 AA.
AC Q6TKD4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Photosystem II protein (Fragment).
GN Name=psbC;
OS Sechium edule.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Cucurbitales; Cucurbitaceae; Sechium.
OX NCBI_TaxID=184140;
RN [1]
RP SEQUENCE FROM N.A.
RA Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY396190; AAR07575.1; -.

```
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 3
Q6TKD5
ID   Q6TKD5      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKD5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Sicyos angulatus.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Sicyos.
OX   NCBI_TaxID=64232;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396189; AAR07574.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 4
Q6TKD6
ID   Q6TKD6      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKD6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Trichosanthes cucumerina.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX   NCBI_TaxID=50543;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396188; AAR07573.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 5
Q6TKD7
ID   Q6TKD7      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKD7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Luffa graveolens.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Luffa.
OX   NCBI_TaxID=252536;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396187; AAR07572.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 6
Q6TKD8
ID   Q6TKD8      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKD8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Luffa echinata.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Luffa.
OX   NCBI_TaxID=252535;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396186; AAR07571.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
```

```
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 7
Q6TKD9
ID   Q6TKD9      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKD9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Cucurbita pepo (Vegetable marrow) (Summer squash).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX   NCBI_TaxID=3663;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396185; AAR07570.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 8
Q6TKE0
ID   Q6TKE0      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKE0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Benincasa hispida (Wax gourd).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Benincasa.
OX   NCBI_TaxID=102211;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396184; AAR07569.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 9
Q6TKE1
ID   Q6TKE1      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKE1;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Marah oreganus.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Marah.
OX   NCBI_TaxID=252534;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396183; AAR07568.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 10
Q6TKE2
ID   Q6TKE2      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKE2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Cyclanthera pedata (Achocha) (Caihua).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Cyclanthera.
OX   NCBI_TaxID=198836;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396182; AAR07567.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 11
Q6TKE3
ID   Q6TKE3      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKE3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
```

```
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Echinocystis lobata.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Echinocystis.
OX   NCBI_TaxID=252533;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396181; AAR07566.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 12
Q6TKE4
ID   Q6TKE4      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKE4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Luffa quinquefida.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Luffa.
OX   NCBI_TaxID=34295;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396180; AAR07565.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 13
Q6TKE5
ID   Q6TKE5      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKE5;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Lagenaria leucantha (Bottle gourd) (Lagenaria siceraria).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Lagenaria.
OX   NCBI_TaxID=3668;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396179; AAR07564.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 14
Q6TKE6
ID   Q6TKE6      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKE6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Citrullus lanatus (Watermelon) (Citrullus vulgaris).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Citrullus.
OX   NCBI_TaxID=3654;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396178; AAR07563.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


RESULT 15
Q6TKE7
ID   Q6TKE7      PRELIMINARY;      PRT;    13 AA.
AC   Q6TKE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Photosystem II protein (Fragment).
GN   Name=psbC;
OS   Citrullus colocynthis.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Cucurbitales; Cucurbitaceae; Citrullus.
OX   NCBI_TaxID=252529;
RN   [1]
```

```
RP   SEQUENCE FROM N.A.
RA   Decker-Walters D.S., Chung S.-M., Staub J.E.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY396177; AAR07562.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1579 MW;  CF0270FBB503C732 CRC64;

  Query Match             33.8%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   62.5%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFL 8
              ||| | |:
Db          2 DFEPVLFM 9


Search completed: February 22, 2005, 09:37:58
Job time : 53.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-9
Perfect score: 77
Sequence: 1 DFEVVTFLKDVLPEF 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	62.5	81.2	14	2	AAW04849	Aaw04849 Internal
2	35	45.5	15	1	AAP82806	Aap82806 Von Wille
3	35	45.5	15	2	AAR40254	Aar40254 von Wille
4	33	42.9	14	2	AAW90258	Aaw90258 Human NBC
5	29	37.7	14	4	AAU69166	Aau69166 Human Ace
6	29	37.7	15	5	AAM47974	Aam47974 Human cyt
7	28	36.4	9	3	AAB06739	Aab06739 Claudin-5
8	28	36.4	9	4	AAB75878	Aab75878 Influenza
9	28	36.4	10	7	ADD23785	Add23785 Breast ca
10	28	36.4	10	8	ADK08674	Adk08674 Human pap
11	28	36.4	14	2	AAW74950	Aaw74950 Human sec
12	28	36.4	14	5	ABG95408	Abg95408 Human nov
13	28	36.4	14	6	ABO34602	Abo34602 Region of
14	28	36.4	14	7	ADI23263	Adi23263 Novel hum
15	28	36.4	14	8	ADH74265	Adh74265 Human sec
16	28	36.4	14	8	ADN65356	Adn65356 HLA bindi
17	28	36.4	15	2	AAW61167	Aaw61167 Ant1 olig
18	27	35.1	9	5	ABG32317	Abg32317 HLA-A1/A2
19	27	35.1	10	4	AAM42905	Aam42905 Mycoplasm
20	27	35.1	13	2	AAR54969	Aar54969 SorhI gra
21	27	35.1	13	2	AAR54970	Aar54970 SorhI gra
22	27	35.1	13	2	AAR72553	Aar72553 Pertussis
23	27	35.1	13	2	AAW29757	Aaw29757 Malassezi
24	27	35.1	13	2	AAY41824	Aay41824 Pertussis
25	27	35.1	13	2	AAW95234	Aaw95234 PT toxin
26	27	35.1	13	2	AAY43371	Aay43371 Pertussis
27	27	35.1	13	3	AAY68373	Aay68373 Pertussis
28	27	35.1	13	4	AAB66247	Aab66247 B pertuss
29	27	35.1	13	4	AAU25795	Aau25795 Breast ca
30	27	35.1	13	5	AAU96084	Aau96084 Peptide s
31	27	35.1	15	2	AAR97502	Aar97502 Mycobacte
32	27	35.1	15	2	AAW04851	Aaw04851 Internal
33	27	35.1	15	2	AAW08922	Aaw08922 Tumour as
34	27	35.1	15	2	AAW59078	Aaw59078 FMDV non-
35	27	35.1	15	2	AAY29716	Aay29716 Mycobacte
36	27	35.1	15	5	ABP30998	Abp30998 Peptide #
37	27	35.1	15	7	ADA08661	Ada08661 Human O77
38	27	35.1	15	7	ADF09004	Adf09004 Secreted
39	27	35.1	15	8	ADN71486	Adn71486 Human 273
40	27	35.1	15	8	ADN70576	Adn70576 Human 273
41	27	35.1	15	8	ADN71729	Adn71729 Human 273
42	27	35.1	15	8	ADN70668	Adn70668 Human 273
43	27	35.1	15	8	ADN71020	Adn71020 Human 273
44	27	35.1	15	8	ADN71487	Adn71487 Human 273
45	27	35.1	15	8	ADN71526	Adn71526 Human 273

ALIGNMENTS

RESULT 1
AAW04849
ID AAW04849 standard; peptide; 14 AA.
XX
AC AAW04849;
XX
DT 16-OCT-2003 (revised)
DT 18-FEB-1997 (first entry)
XX
DE Internal fragment of adenovirus type 12 ORF.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Human adenovirus type 12.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 2; Page 43; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with

```
CC   the autoimmune disease. This peptide is an internal peptide of the
CC   adenovirus type 12 ORF and is implicated as a foreign epitope involved in
CC   the aetiology or in remissions of multiple sclerosis. It has been shown
CC   capable of inducing the proliferation of autoreactive T-cell clones
CC   isolated from multiple sclerosis patients. (Updated on 16-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 14 AA;

  Query Match             81.2%;  Score 62.5;  DB 2;  Length 14;
  Best Local Similarity   93.3%;  Pred. No. 0.0012;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          1 DFEVVTFLKDVLPEF 15
              ||||||||||| |||
Db          1 DFEVVTFLKDV-PEF 14


RESULT 2
AAP82806
ID   AAP82806 standard; protein; 15 AA.
XX
AC   AAP82806;
XX
DT   25-MAR-2003  (revised)
DT   12-DEC-1990  (first entry)
XX
DE   Von Willebrand Factor-inhibiting peptide #46.
XX
KW   Von Willebrand Factor; thrombosis; platelet binding inhibitor.
XX
OS   Synthetic.
XX
PN   AU8773715-A.
XX
PD   03-DEC-1987.
XX
PF   01-JUN-1987;   87AU-00073715.
XX
PR   30-MAY-1986;   86US-00869188.
XX
PA   (SCRI ) SCRIPPS CLINIC & RE.
XX
DR   WPI; 1988-021781/04.
XX
PT   New peptide fragments of Von Willebrand factor - inhibit binding of Von
PT   Willebrand factor to platelets, heparin and collagen, used for treating
PT   thrombosis.
XX
PS   Claim 2; Page 22; 29pp; English.
XX
CC   This 15 residue inhibitory peptide is based on part of the sequence of a
CC   52/48KD polypeptide fragment from Von Willebrand factor. The polypeptide
CC   corresponds to residues 449-729 of vWF. Other 15 residue peptide based on
CC   other regions of the same fragment of vWF are also found to inhibit
CC   binding to platelets and heparin. A more potent inhibitor could be
CC   produced by linking two or more of these peptides. On its own this
CC   peptide reduces vWF binding to platelets to 20% (c.f. 52/48kD fragment
CC   reduces binding to 0%). See also AAP82761-P82805 and AAP82807-P82810.
CC   (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 15 AA;

  Query Match             45.5%;  Score 35;  DB 1;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 60;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 EVVTFLKDVLPE 14
              |:|::| |: ||
Db          1 EIVSYLCDLAPE 12


RESULT 3
AAR40254
ID   AAR40254 standard; protein; 15 AA.
XX
AC   AAR40254;
XX
DT   25-MAR-2003  (revised)
DT   15-FEB-1994  (first entry)
XX
DE   von Willebrand Factor/platelet binding inhibitor peptide.
XX
KW   vWF; pMMB3; expression; binding inhibition; inhibiting; prevention;
KW   treatment; cardiovascular disorders; thrombosis.
XX
OS   Synthetic.
XX
PN   US5238919-A.
XX
PD   24-AUG-1993.
XX
PF   07-MAY-1990;   90US-00519606.
XX
PR   30-MAY-1986;   86US-00869188.
PR   04-NOV-1988;   88US-00270488.
XX
PA   (SCRI ) SCRIPPS CLINIC & RES FOUND.
XX
PI   Zimmerman TS,  Fujimura Y,  Houghten RA,  Ruggeri ZM;
XX
DR   WPI; 1993-280681/35.
XX
PT   Use of Von=Willebrand Factor fragment peptide(s) - for inhibiting binding
PT   of factor to platelets, heparin and collagen, partic. for treating
PT   thrombosis.
XX
PS   Claim 3; Page 26; 26pp; English.
XX
CC   The sequence is that of a peptide which is capable of inhibiting the
CC   binding of von Willebrand factor (vWF) to platelets. As such it can be
CC   used for the prevention and treatment of cardiovascular disorders such as
CC   thrombosis. (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ   Sequence 15 AA;

  Query Match             45.5%;  Score 35;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 60;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 EVVTFLKDVLPE 14
              |:|::| |: ||
Db          1 EIVSYLCDLAPE 12


RESULT 4
AAW90258
ID   AAW90258 standard; peptide; 14 AA.
XX
AC   AAW90258;
XX
DT   27-APR-1999  (first entry)
XX
DE   Human NBC protein fragment #1.
XX
KW   Anion exchange; NBC; sodium bicarbonate transporter family; DIDS motif;
KW   treatment; water retention; blood pressure; acidosis; inflammation;
KW   cell proliferation; cancer; sperm activation; inactivation; epilepsy;
KW   hydroencephaly; glaucoma; colitis; pH regulation; immunoassay.
XX
OS   Homo sapiens.
XX
PN   WO9853067-A1.
XX
PD   26-NOV-1998.
XX
```

PF 20-MAY-1998; 98WO-US010297.
XX
PR 20-MAY-1997; 97US-0047131P.
XX
PA (UYYA) UNIV YALE.
XX
PI Bevensee MO, Schmitt BM, Romero MF, Boron WF, Biemesderfer D;
PI Davis BA, Sussman CR, Choi I, Aalkjaer C, Grichtchenko II;
XX
DR WPI; 1999-059743/05.
XX
PT New nucleic acid molecules encoding proteins of the Sodium Bicarbonate
PT Cotransporter (NBC) family - useful for identifying agents that agonise
PT or antagonise NBC activity and treating disorders mediated by NBC.
XX
PS Example 3; Page 44; 138pp; English.
XX
CC This peptide fragment represents an anion exchange DIDS binding motif
CC which is found in a novel sodium bicarbonate transporter (NBC). The novel
CC NBC proteins and nucleic acid sequences may be used to treat pathological
CC processes including water retention, increased blood pressure, chronic
CC respiratory and metabolic acidosis, inflammation, cell proliferation,
CC cancer, sperm activation/inactivation, hydroencephaly, epilepsy, glaucoma
CC and colitis. Members of the NBC family of proteins can be used (i) as a
CC target to identify agents that block or stimulate NBC mediated pH
CC regulation, (ii) as a target or bait to identify and isolate binding
CC partners that bind NBC, (iii) in methods to identify agents that block or
CC stimulate activity of an NBC protein, and (iv) as a target to assay for
CC NBC-mediated activity. Anti-NBC antibodies are also useful as modulators
CC of NBC activity, useful in the immunoassays for detecting NBC
CC expression/activity and for purifying an NBC protein
XX
SQ Sequence 14 AA;

Query Match 42.9%; Score 33; DB 2; Length 14;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 FLKDVLPE 14
|| ||:||
Db 1 FLDDVIPE 8

RESULT 5
AAU69166
ID AAU69166 standard; peptide; 14 AA.
XX
AC AAU69166;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human Acetylcholine receptor alpha subunit peptide AChR 99-112.
XX
KW Human; epitope; autoimmune disease; myasthenia gravis;
KW Human leukocyte antigen; acetylcholine receptor; HLA DR3; HLA DR2; AChR;
KW antigen; immunosuppressive; major histocompatibility complex; MHC.
XX
OS Homo sapiens.
XX
PN WO200174848-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US010450.
XX
PR 31-MAR-2000; 2000US-0193745P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Deshpande S, Spack E, Wehner N, Arimilli S;
XX
DR WPI; 2001-648547/74.
XX
PT Peptide epitopes of the acetylcholine receptor target helper T cells
PT recognize an antigen in association with an MHC component and are useful
PT to treat autoimmune disease particularly myasthenia gravis.
XX
PS Example 1; Fig 2; 46pp; English.
XX
CC The invention relates to a composition comprising an isolated
CC acetylcholine receptor (AChR) oligopeptide of about 12 to 20 amino acids.
CC The peptides form a set of 69 overlapping antigenic peptide epitopes
CC which show various affinities for human leukocyte antigens (HLA) HLA-DR2
CC and DR3. Also included is a composition comprising an antigenic peptide
CC and a Major histocompatibility complex (MHC) component having an
CC antigenic binding site, where binding of the peptide to the binding site
CC induces non-responsiveness in a target T cell in a mammal, where the MHC
CC component is an MHC class II component. Peptides with affinity for HLA-
CC DR2 and DR-3 from proteins other than AChR are also included. The
CC composition is used to treat myasthenia gravis and other autoimmune
CC diseases. The present sequence is an antigenic peptide epitope
XX
SQ Sequence 14 AA;

Query Match 37.7%; Score 29; DB 4; Length 14;
Best Local Similarity 50.0%; Pred. No. 5.8e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 DFEVVTFLKDVL 12
|| :| | | :|
Db 1 DFAIVKFTKVLL 12

RESULT 6
AAM47974
ID AAM47974 standard; peptide; 15 AA.
XX
AC AAM47974;
XX
DT 06-MAR-2002 (first entry)
XX
DE Human cytokinesis regulatory protein 55 peptide SEQ ID NO 7.
XX
KW Human; cytokinesis regulatory protein 55; cancer; haemopathy; HIV;
KW human immunodeficiency virus; infection.
XX
OS Homo sapiens.
XX
PN CN1315368-A.
XX
PD 03-OCT-2001.
XX
PF 27-MAR-2000; 2000CN-00115139.
XX
PR 27-MAR-2000; 2000CN-00115139.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-056314/08.
XX
PT Polypeptide-human cytokinesis regulating protein 55 and polynucleotide
PT for coding it.
XX
PS Example 6; Page 19 (Disclosure); 32pp; Chinese.
XX
CC The invention relates to human cytokinesis regulatory protein 55, the
CC polynucleotide encoding it, the process for preparing the polypeptide by
CC DNA recombinant technology, the application of the polypeptide in
CC treating several diseases such as cancer, haemopathy, HIV infection. The
CC present sequence is that of a peptide, useful to the invention
XX
SQ Sequence 15 AA;

Query Match 37.7%; Score 29; DB 5; Length 15;

```
  Best Local Similarity   50.0%;  Pred. No. 6.2e+02;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 LKDVLPEF 15
              :|:::|||
Db          4 VKELIPEF 11


RESULT 7
AAB06739
ID   AAB06739 standard; peptide; 9 AA.
XX
AC   AAB06739;
XX
DT   28-SEP-2000  (first entry)
XX
DE   Claudin-5 cyclic cell adhesion recognition sequence SEQ ID NO: 310.
XX
KW   Claudin-5 modulating agent; cell adhesion recognition sequence;
KW   CAR sequence; autoimmune disease; inflammatory disease; cancer;
KW   graft rejection; cyclic.
XX
OS   Mammalia.
XX
PN   WO200026360-A1.
XX
PD   11-MAY-2000.
XX
PF   03-NOV-1999;   99WO-CA001029.
XX
PR   03-NOV-1998;   98US-00185908.
PR   30-MAR-1999;   99US-00282029.
XX
PA   (ADHE-) ADHEREX TECHNOLOGIES INC.
XX
PI   Blaschuck OW,  Symonds JM,  Gour BJ;
XX
DR   WPI; 2000-365610/31.
XX
PT   Antibody modulation of claudin-mediated cell adhesion for increasing
PT   vasopermiability, for delivering drugs to tumors and the nervous system
PT   and across the skin.
XX
PS   Claim 67; Page 102; 121pp; English.
XX
CC   The present invention relates to the use of peptides as claudin-mediated
CC   cell adhesion modulators. The claudin-5 group of proteins are cadherins,
CC   which are membrane glycoproteins involved in cell adhesion. In some
CC   situations, cell adhesion occurs at abnormal levels, and these peptides
CC   can be used to modulate these levels, and thus treat autoimmune diseases,
CC   inflammatory diseases and cancer, and aid wound healing and implant
CC   adhesion. In addition, they can also be used to facilitate drug delivery
CC   to the desired target site. The present sequence has a cyclic
CC   conformation
XX
SQ   Sequence 9 AA;

  Query Match             36.4%;  Score 28;  DB 3;  Length 9;
  Best Local Similarity   55.6%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DFEVVTFLK 9
              |::|  |||
Db          1 DWQVTAFLK 9


RESULT 8
AAB75878
ID   AAB75878 standard; peptide; 9 AA.
XX
AC   AAB75878;
XX
DT   10-APR-2001  (first entry)
XX
DE   Influenza virus HLA-A2 binding peptide.
XX
KW   Human leukocyte antigen; HLA; major histocompatibility complex; MHC;
KW   cytotoxic T lymphocyte; CTL; human class I MHC; immunogenic;
KW   HLA binding peptide; immune response; glycoprotein; cytostatic; virucide;
KW   hepatotropic; antiinflammatory; anti-HIV; vaccine;
KW   human immunodeficiency virus; protozoacide; viral infection; cancer;
KW   prostate cancer; hepatitis B; hepatitis C; human papilloma virus; HPV;
KW   cytomegalovirus; CMV; acquired immunodeficiency syndrome; AIDS;
KW   renal carcinoma; cervical carcinoma; lymphoma; malaria;
KW   condyloma acuminatum.
XX
OS   Influenza virus.
XX
PN   WO200100225-A1.
XX
PD   04-JAN-2001.
XX
PF   28-JUN-2000; 2000WO-US017842.
XX
PR   29-JUN-1999;   99US-0141422P.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 2001-112389/12.
XX
PT   Composition comprising human leukocyte antigen binding peptide which
PT   comprises isolated, prepared epitope useful for treating viral infections
PT   such as acquired immunodeficiency syndrome, and cancer.
XX
PS   Claim 1; Page 42; 58pp; English.
XX
CC   The present invention describes a composition (I) which comprises at
CC   least one human leukocyte antigen (HLA) binding peptide comprising an
CC   isolated, prepared epitope comprising one of 547 8-11 residue amino acid
CC   sequences (S1), given in AAB75803 to AAB76349. (I) has cytostatic,
CC   virucide, hepatotropic, antiinflammatory, anti-HIV (human
CC   immunodeficiency virus) and protozoacide activities, which can be used in
CC   vaccine production and is an inducer of cytotoxic T-cell response. (I) is
CC   useful for inducing a cytotoxic T cell response against a preselected
CC   antigen in a patient expressing a specific major histocompatibility
CC   complex (MHC) class I allele, by contacting cytotoxic T cells (CTLs) from
CC   the patient with (I). (I) is useful as a vaccine to treat and/or prevent
CC   viral infection and cancer such as prostate cancer, hepatitis B,
CC   hepatitis C, human papilloma virus (HPV) infection, cytomegalovirus
CC   (CMV), acquired immunodeficiency syndrome (AIDS), renal carcinoma,
CC   cervical carcinoma, lymphoma, malaria, and condyloma acuminatum
XX
SQ   Sequence 9 AA;

  Query Match             36.4%;  Score 28;  DB 4;  Length 9;
  Best Local Similarity   62.5%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 VVTFLKDV 11
              :: |||||
Db          2 LIDFLKDV 9


RESULT 9
ADD23785
ID   ADD23785 standard; peptide; 10 AA.
XX
AC   ADD23785;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Breast cancer membrane protein (BCMP) peptide SEQ ID NO:368.
XX
KW   breast cancer; screening; diagnosis; breast cancer therapy;
```

KW breast cancer membrane protein; BCMP; cytostatic; vaccine; human.
XX
OS Homo sapiens.
XX
PN WO2003087831-A2.
XX
PD 23-OCT-2003.
XX
PF 10-APR-2003; 2003WO-GB001559.
XX
PR 11-APR-2002; 2002GB-00008331.
XX
PA (OXFO-) OXFORD GLYCOSCIENCES UK LTD.
XX
PI Hudson LJ, Stamps AC, Terrett JA;
XX
DR WPI; 2003-845381/78.
XX
PT Screening, diagnosing and/or treating breast cancer by detecting a change
PT in expression or activity of a breast cancer membrane protein (BCMP)
PT polypeptide or encoding nucleic acid molecule.
XX
PS Claim 1; SEQ ID NO 368; 81pp; English.
XX
CC The present invention describes a method of screening for and/or
CC diagnosing breast cancer in a subject, and/or monitoring the
CC effectiveness of breast cancer therapy. The method comprises detecting
CC and/or quantifying in a biological sample obtained from the subject a
CC breast cancer membrane protein (BCMP) polypeptide and a nucleic acid
CC molecule. Also described: (1) an antibody, its functionally-active
CC fragment, derivative or analogue, that specifically binds to one or more
CC of the BCMP polypeptide; (2) a diagnostic kit comprising a capture
CC reagent specific for an BCMP polypeptide, reagents and instructions for
CC use; (3) a method for screening for anti-breast cancer agents that
CC interact with the BCMP polypeptide, comprising contacting the polypeptide
CC with a candidate agent, and determining whether or not the candidate
CC agent interacts with the polypeptide; (4) a method for screening for anti
CC -breast cancer agents that modulate the expression or activity of an BCMP
CC polypeptide or the nucleic acid molecule cited above, comprising
CC comparing the expression or activity of the polypeptide or nucleic acid
CC molecule, in the presence and absence of a candidate agent or in the
CC presence of a control agent, and determining whether the candidate agent
CC causes the expression or activity of the polypeptide or nucleic acid
CC molecule to change; and (5) an agent identified by the method of (3) or
CC (4), which interacts with the polypeptide or causes the expression or
CC activity of the polypeptide, or the expression of the nucleic acid
CC molecule to change. BCMPs have cytostatic activities, and can be used in
CC vaccines. The BCMP polypeptide, nucleic acid molecule, antibody, agent or
CC their derivatives, are useful in the manufacture of a medicament for the
CC treatment of breast cancer, where the composition is a vaccine. The
CC present sequence represents a BCMP peptide which is used in the
CC exemplification of the present invention.
XX
SQ Sequence 10 AA;

```
  Query Match             36.4%;  Score 28;  DB 7;  Length 10;
  Best Local Similarity   62.5%;  Pred. No. 6e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 LKDVLPEF 15
              |:|  |||
Db          1 LEDYFPEF 8
```

RESULT 10
ADK08674
ID ADK08674 standard; peptide; 10 AA.
XX
AC ADK08674;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human papillomavirus peptide #729.
XX
KW pathogenic virus; alternative reading frame; antigenic determinant;
KW virucide; vaccine; therapeutic agent; infection; HPV.
XX
OS Human papillomavirus.
XX
PN WO2004011650-A2.
XX
PD 05-FEB-2004.
XX
PF 24-JUL-2003; 2003WO-EP008112.
XX
PR 24-JUL-2002; 2002AT-00001124.
PR 11-JUL-2003; 2003EP-00450171.
XX
PA (INTE-) INTERCELL AG.
XX
PI Mattner F, Schmidt W, Habel A;
XX
DR WPI; 2004-169243/16.
XX
PT New polypeptide encoded by an alternative reading frame of a pathogenic
PT virus comprising an antigenic determinant, useful for treating or
PT preventing an infection with the pathogenic virus.
XX
PS Claim 18; Page 181; 220pp; English.
XX
CC This invention relates to a novel polypeptide encoded by an alternative
CC reading frame of a pathogenic virus, where the polypeptide starts with a
CC methionine amino acid residue, which comprises an antigenic determinant
CC and more than 7 amino acid residues. The invention may be useful for the
CC production of compounds with a virucide activity or the development of a
CC vaccine. The polypeptide or its fragments may be useful as a therapeutic
CC agent. It is also useful for the manufacture of a medicament for treating
CC or preventing an infection with the pathogenic virus. The present
CC sequence is that of a human papillomavirus (HPV) epitope peptide of the
CC invention.
XX
SQ Sequence 10 AA;

```
  Query Match             36.4%;  Score 28;  DB 8;  Length 10;
  Best Local Similarity   62.5%;  Pred. No. 6e+02;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 VTFLKDVL 12
              : |||||:
Db          1 ILFLKDVV 8
```

RESULT 11
AAW74950
ID AAW74950 standard; protein; 14 AA.
XX
AC AAW74950;
XX
DT 19-JAN-1999 (first entry)
XX
DE Human secreted protein encoded by gene 66 clone HTOCD52.
XX
KW Human; secreted protein; testis; tumour; foetal brain tissue;
KW fusion protein; cancer; central nervous system; seizure; diagnosis;
KW neurodegenerative disease.
XX
OS Homo sapiens.
XX
PN WO9839448-A2.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US004493.
XX
PR 07-MAR-1997; 97US-0038621P.
PR 07-MAR-1997; 97US-0040161P.

PR 07-MAR-1997; 97US-0040162P.
PR 07-MAR-1997; 97US-0040163P.
PR 07-MAR-1997; 97US-0040333P.
PR 07-MAR-1997; 97US-0040334P.
PR 07-MAR-1997; 97US-0040336P.
PR 07-MAR-1997; 97US-0040626P.
PR 11-APR-1997; 97US-0043311P.
PR 11-APR-1997; 97US-0043312P.
PR 11-APR-1997; 97US-0043313P.
PR 11-APR-1997; 97US-0043314P.
PR 11-APR-1997; 97US-0043315P.
PR 11-APR-1997; 97US-0043568P.
PR 11-APR-1997; 97US-0043569P.
PR 11-APR-1997; 97US-0043576P.
PR 11-APR-1997; 97US-0043578P.
PR 11-APR-1997; 97US-0043580P.
PR 11-APR-1997; 97US-0043669P.
PR 11-APR-1997; 97US-0043670P.
PR 11-APR-1997; 97US-0043671P.
PR 11-APR-1997; 97US-0043672P.
PR 11-APR-1997; 97US-0043674P.
PR 23-MAY-1997; 97US-0047492P.
PR 23-MAY-1997; 97US-0047500P.
PR 23-MAY-1997; 97US-0047501P.
PR 23-MAY-1997; 97US-0047502P.
PR 23-MAY-1997; 97US-0047503P.
PR 23-MAY-1997; 97US-0047581P.
PR 23-MAY-1997; 97US-0047582P.
PR 23-MAY-1997; 97US-0047583P.
PR 23-MAY-1997; 97US-0047584P.
PR 23-MAY-1997; 97US-0047585P.
PR 23-MAY-1997; 97US-0047586P.
PR 23-MAY-1997; 97US-0047587P.
PR 23-MAY-1997; 97US-0047588P.
PR 23-MAY-1997; 97US-0047589P.
PR 23-MAY-1997; 97US-0047590P.
PR 23-MAY-1997; 97US-0047592P.
PR 23-MAY-1997; 97US-0047593P.
PR 23-MAY-1997; 97US-0047594P.
PR 23-MAY-1997; 97US-0047595P.
PR 23-MAY-1997; 97US-0047596P.
PR 23-MAY-1997; 97US-0047597P.
PR 23-MAY-1997; 97US-0047598P.
PR 23-MAY-1997; 97US-0047599P.
PR 23-MAY-1997; 97US-0047600P.
PR 23-MAY-1997; 97US-0047601P.
PR 23-MAY-1997; 97US-0047612P.
PR 23-MAY-1997; 97US-0047613P.
PR 23-MAY-1997; 97US-0047614P.
PR 23-MAY-1997; 97US-0047615P.
PR 23-MAY-1997; 97US-0047617P.
PR 23-MAY-1997; 97US-0047618P.
PR 23-MAY-1997; 97US-0047632P.
PR 23-MAY-1997; 97US-0047633P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048974P.
PR 13-JUN-1997; 97US-0049610P.
PR 08-JUL-1997; 97US-0051926P.
PR 16-JUL-1997; 97US-0052874P.
PR 18-AUG-1997; 97US-0055724P.
PR 22-AUG-1997; 97US-0056630P.
PR 22-AUG-1997; 97US-0056631P.
PR 22-AUG-1997; 97US-0056632P.
PR 22-AUG-1997; 97US-0056636P.
PR 22-AUG-1997; 97US-0056637P.
PR 22-AUG-1997; 97US-0056662P.
PR 22-AUG-1997; 97US-0056664P.
PR 22-AUG-1997; 97US-0056845P.
PR 22-AUG-1997; 97US-0056862P.
PR 22-AUG-1997; 97US-0056864P.
PR 22-AUG-1997; 97US-0056872P.
PR 22-AUG-1997; 97US-0056874P.
PR 22-AUG-1997; 97US-0056875P.
PR 22-AUG-1997; 97US-0056876P.
PR 22-AUG-1997; 97US-0056877P.
PR 22-AUG-1997; 97US-0056878P.
PR 22-AUG-1997; 97US-0056879P.
PR 22-AUG-1997; 97US-0056880P.
PR 22-AUG-1997; 97US-0056881P.
PR 22-AUG-1997; 97US-0056882P.
PR 22-AUG-1997; 97US-0056884P.
PR 22-AUG-1997; 97US-0056886P.
PR 22-AUG-1997; 97US-0056887P.
PR 22-AUG-1997; 97US-0056888P.
PR 22-AUG-1997; 97US-0056889P.
PR 22-AUG-1997; 97US-0056892P.
PR 22-AUG-1997; 97US-0056893P.
PR 22-AUG-1997; 97US-0056894P.
PR 22-AUG-1997; 97US-0056903P.
PR 22-AUG-1997; 97US-0056908P.
PR 22-AUG-1997; 97US-0056909P.
PR 22-AUG-1997; 97US-0056910P.
PR 22-AUG-1997; 97US-0056911P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057669P.
PR 05-SEP-1997; 97US-0057761P.
PR 12-SEP-1997; 97US-0058785P.
PR 02-OCT-1997; 97US-0061060P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Rosen CA, Fischer CL, Soppet DR, Carter KC;
PI Bednarik DP, Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM;
PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA;
PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;
XX
DR WPI; 1998-506364/43.
DR N-PSDB; AAV59735.
XX
PT New isolated human genes and the secreted polypeptide(s) they encode -
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
PS Claim 1; Page 670; 721pp; English.
XX
CC This sequence represents a secreted human protein encoded by the nucleic
CC acid molecule designated Gene 66 from the human cDNA clone HTOCD52
CC (deposited as clone ATCC 97900 and ATCC 209046). The gene can be used to
CC generate fusion proteins by linking to the gene to a human immunoglobulin
CC Fc portion (e.g. AAV59502) for increasing the stability of the fused
CC protein as compared to the human protein only. The invention relates to
CC 186 novel genes and their fragments (nucleic acid sequences: AAV59511-
CC V59812; amino acid sequences AAW74731-W75026) which are useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. Also, pathological conditions can be diagnosed by
CC determining the amount of the new polypeptides in a sample or by
CC determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 186 polynucleotides, based on
CC which tissues they are most highly expressed in (see AAV59511 for
CC described uses)
XX
SQ Sequence 14 AA;

Query Match 36.4%; Score 28; DB 2; Length 14;
Best Local Similarity 36.4%; Pred. No. 8.5e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FEVVTFLKDVL 12
|: ::: ||:|
Db 2 FDFLSYFKDLL 12

RESULT 12
ABG95408
ID ABG95408 standard; protein; 14 AA.
XX

AC ABG95408;
XX
DT 15-JAN-2003 (first entry)
XX
DE Human novel secreted protein #229.
XX
KW Human; secreted protein; autoimmune disease; chemotaxis;
KW rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;
KW liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;
KW cerebrovascular disorder; cerebral ischaemia; angiogenesis; sunburn;
KW nervous system disorders; Alzheimer's disease; infection;
KW ocular disorder; corneal infection; wound healing; tissue regeneration;
KW epithelial cell proliferation; organ transplantation; food additive;
KW preservative; nutritional.
XX
OS Homo sapiens.
XX
PN US6420526-B1.
XX
PD 16-JUL-2002.
XX
PF 08-SEP-1998; 98US-00149476.
XX
PR 07-MAR-1997; 97US-0038621P.
PR 07-MAR-1997; 97US-0040161P.
PR 07-MAR-1997; 97US-0040162P.
PR 07-MAR-1997; 97US-0040163P.
PR 07-MAR-1997; 97US-0040333P.
PR 07-MAR-1997; 97US-0040334P.
PR 07-MAR-1997; 97US-0040336P.
PR 07-MAR-1997; 97US-0040626P.
PR 11-APR-1997; 97US-0043311P.
PR 11-APR-1997; 97US-0043312P.
PR 11-APR-1997; 97US-0043313P.
PR 11-APR-1997; 97US-0043314P.
PR 11-APR-1997; 97US-0043315P.
PR 11-APR-1997; 97US-0043568P.
PR 11-APR-1997; 97US-0043569P.
PR 11-APR-1997; 97US-0043576P.
PR 11-APR-1997; 97US-0043578P.
PR 11-APR-1997; 97US-0043580P.
PR 11-APR-1997; 97US-0043669P.
PR 11-APR-1997; 97US-0043670P.
PR 11-APR-1997; 97US-0043671P.
PR 11-APR-1997; 97US-0043672P.
PR 11-APR-1997; 97US-0043674P.
PR 23-MAY-1997; 97US-0047492P.
PR 23-MAY-1997; 97US-0047500P.
PR 23-MAY-1997; 97US-0047501P.
PR 23-MAY-1997; 97US-0047502P.
PR 23-MAY-1997; 97US-0047503P.
PR 23-MAY-1997; 97US-0047581P.
PR 23-MAY-1997; 97US-0047582P.
PR 23-MAY-1997; 97US-0047583P.
PR 23-MAY-1997; 97US-0047584P.
PR 23-MAY-1997; 97US-0047585P.
PR 23-MAY-1997; 97US-0047586P.
PR 23-MAY-1997; 97US-0047587P.
PR 23-MAY-1997; 97US-0047588P.
PR 23-MAY-1997; 97US-0047589P.
PR 23-MAY-1997; 97US-0047590P.
PR 23-MAY-1997; 97US-0047592P.
PR 23-MAY-1997; 97US-0047593P.
PR 23-MAY-1997; 97US-0047594P.
PR 23-MAY-1997; 97US-0047595P.
PR 23-MAY-1997; 97US-0047596P.
PR 23-MAY-1997; 97US-0047597P.
PR 23-MAY-1997; 97US-0047598P.
PR 23-MAY-1997; 97US-0047599P.
PR 23-MAY-1997; 97US-0047600P.
PR 23-MAY-1997; 97US-0047601P.
PR 23-MAY-1997; 97US-0047612P.
PR 23-MAY-1997; 97US-0047613P.
PR 23-MAY-1997; 97US-0047614P.
PR 23-MAY-1997; 97US-0047615P.
PR 23-MAY-1997; 97US-0047617P.
PR 23-MAY-1997; 97US-0047618P.
PR 23-MAY-1997; 97US-0047632P.
PR 23-MAY-1997; 97US-0047633P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048974P.
PR 13-JUN-1997; 97US-0049610P.
PR 08-JUL-1997; 97US-0051926P.
PR 16-JUL-1997; 97US-0052874P.
PR 18-AUG-1997; 97US-0055724P.
PR 22-AUG-1997; 97US-0056630P.
PR 22-AUG-1997; 97US-0056631P.
PR 22-AUG-1997; 97US-0056632P.
PR 22-AUG-1997; 97US-0056636P.
PR 22-AUG-1997; 97US-0056637P.
PR 22-AUG-1997; 97US-0056662P.
PR 22-AUG-1997; 97US-0056664P.
PR 22-AUG-1997; 97US-0056845P.
PR 22-AUG-1997; 97US-0056862P.
PR 22-AUG-1997; 97US-0056864P.
PR 22-AUG-1997; 97US-0056872P.
PR 22-AUG-1997; 97US-0056874P.
PR 22-AUG-1997; 97US-0056875P.
PR 22-AUG-1997; 97US-0056876P.
PR 22-AUG-1997; 97US-0056877P.
PR 22-AUG-1997; 97US-0056878P.
PR 22-AUG-1997; 97US-0056879P.
PR 22-AUG-1997; 97US-0056880P.
PR 22-AUG-1997; 97US-0056881P.
PR 22-AUG-1997; 97US-0056882P.
PR 22-AUG-1997; 97US-0056884P.
PR 22-AUG-1997; 97US-0056886P.
PR 22-AUG-1997; 97US-0056887P.
PR 22-AUG-1997; 97US-0056888P.
PR 22-AUG-1997; 97US-0056889P.
PR 22-AUG-1997; 97US-0056892P.
PR 22-AUG-1997; 97US-0056893P.
PR 22-AUG-1997; 97US-0056894P.
PR 22-AUG-1997; 97US-0056903P.
PR 22-AUG-1997; 97US-0056908P.
PR 22-AUG-1997; 97US-0056909P.
PR 22-AUG-1997; 97US-0056910P.
PR 22-AUG-1997; 97US-0056911P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057669P.
PR 05-SEP-1997; 97US-0057761P.
PR 12-SEP-1997; 97US-0058785P.
PR 02-OCT-1997; 97US-0061060P.
PR 06-MAR-1998; 98WO-US004493.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Rosen CA, Fischer CL, Soppet DP, Carter KC;
PI Bednarik DR, Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM;
PI Ferrie AM, Duan R, Hu J, Florence KA, Olsen HS, Ebner R, Brewer LA;
PI Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;
XX
DR WPI; 2002-634796/68.
DR N-PSDB; ABS73726.
XX
PT New isolated human secreted protein for diagnosing, preventing, treating
PT or ameliorating medical conditions and used as a food additive or
PT preservative.
XX
PS Example 1; SEQ ID NO 548; 129pp; English.
XX
CC The invention relates to an isolated protein that is one of 186 human
CC secreted proteins, given in the specification, encoded by one of 309 cDNA
CC sequences also given in the specification. The protein is used in a
CC pharmaceutical composition used to prevent, treat or ameliorate a medical
CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,

CC chickens or sheep. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC e.g. corneal infection. The polypeptides can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
CC carbohydrate, vitamins, minerals, cofactors and other nutritional
CC components. The present sequence represents one of the novel human
CC secreted proteins of the invention. Note: This sequence did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=6420526B1
XX
SQ Sequence 14 AA;

Query Match 36.4%; Score 28; DB 5; Length 14;
Best Local Similarity 36.4%; Pred. No. 8.5e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 FEVVTFLKDVL 12
              |: ::: ||:|
Db          2 FDFLSYFKDLL 12
```

RESULT 13
ABO34602
ID ABO34602 standard; protein; 14 AA.
XX
AC ABO34602;
XX
DT 22-SEP-2003 (first entry)
XX
DE Region of human secreted protein encoded by cDNA sequence #229.
XX
KW Human; secreted protein; hyperproliferative disorder; leukaemia;
KW breast cancer; wound; reproductive disorder; blood-related disorder;
KW haemophilia; thrombocytopaenia; immunodeficiency; thymic hypoplasia;
KW Wiskott-Aldrich syndrome; autoimmune disorder; multiple sclerosis;
KW graft-versus-host disease; Hashimoto's thyroiditis; allergy; asthma;
KW viral infection; bacterial infection; fungal infection; AIDS; sepsis;
KW renal disorder; kidney failure; cardiovascular disorder; cytostatic;
KW angina pectoris; cerebral ischaemia; congenital heart defect;
KW respiratory disorder; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; inflammation; Crohn's disease; vulnerary;
KW immunosuppressive; antibacterial; haemostatic; thrombolytic;
KW anticoagulant; neuroprotective; thyromimetic; antiallergic;
KW antiasthmatic; virucide; fungicide; anti-HIV; nephrotropic; antianginal;
KW cerebroprotective; cardiant; nootropic; antiparkinsonian;
KW antiinflammatory.
XX
OS Homo sapiens.
XX
PN US2003049618-A1.
XX
PD 13-MAR-2003.
XX
PF 16-MAR-2001; 2001US-00809391.
XX
PR 07-MAR-1997; 97US-0038621P.
PR 07-MAR-1997; 97US-0040162P.
PR 07-MAR-1997; 97US-0040163P.
PR 07-MAR-1997; 97US-0040333P.
PR 07-MAR-1997; 97US-0040334P.
PR 07-MAR-1997; 97US-0040336P.
PR 07-MAR-1997; 97US-0040626P.
PR 11-APR-1997; 97US-0043311P.
PR 11-APR-1997; 97US-0043312P.
PR 11-APR-1997; 97US-0043313P.
PR 11-APR-1997; 97US-0043314P.
PR 11-APR-1997; 97US-0043315P.
PR 11-APR-1997; 97US-0043568P.
PR 11-APR-1997; 97US-0043569P.
PR 11-APR-1997; 97US-0043576P.
PR 11-APR-1997; 97US-0043578P.
PR 11-APR-1997; 97US-0043580P.
PR 11-APR-1997; 97US-0043669P.
PR 11-APR-1997; 97US-0043670P.
PR 11-APR-1997; 97US-0043671P.
PR 11-APR-1997; 97US-0043672P.
PR 11-APR-1997; 97US-0043674P.
PR 23-MAY-1997; 97US-0047492P.
PR 23-MAY-1997; 97US-0047500P.
PR 23-MAY-1997; 97US-0047501P.
PR 23-MAY-1997; 97US-0047502P.
PR 23-MAY-1997; 97US-0047503P.
PR 23-MAY-1997; 97US-0047581P.
PR 23-MAY-1997; 97US-0047582P.
PR 23-MAY-1997; 97US-0047583P.
PR 23-MAY-1997; 97US-0047584P.
PR 23-MAY-1997; 97US-0047585P.
PR 23-MAY-1997; 97US-0047586P.
PR 23-MAY-1997; 97US-0047587P.
PR 23-MAY-1997; 97US-0047588P.
PR 23-MAY-1997; 97US-0047589P.
PR 23-MAY-1997; 97US-0047590P.
PR 23-MAY-1997; 97US-0047592P.
PR 23-MAY-1997; 97US-0047593P.
PR 23-MAY-1997; 97US-0047594P.
PR 23-MAY-1997; 97US-0047595P.
PR 23-MAY-1997; 97US-0047596P.
PR 23-MAY-1997; 97US-0047597P.
PR 23-MAY-1997; 97US-0047598P.
PR 23-MAY-1997; 97US-0047599P.
PR 23-MAY-1997; 97US-0047600P.
PR 23-MAY-1997; 97US-0047601P.
PR 23-MAY-1997; 97US-0047612P.
PR 23-MAY-1997; 97US-0047613P.
PR 23-MAY-1997; 97US-0047614P.
PR 23-MAY-1997; 97US-0047615P.
PR 23-MAY-1997; 97US-0047617P.
PR 23-MAY-1997; 97US-0047618P.
PR 23-MAY-1997; 97US-0047632P.
PR 23-MAY-1997; 97US-0047633P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048974P.
PR 13-JUN-1997; 97US-0049610P.
PR 08-JUL-1997; 97US-0051926P.
PR 16-JUL-1997; 97US-0052874P.
PR 18-AUG-1997; 97US-0055724P.
PR 22-AUG-1997; 97US-0056630P.
PR 22-AUG-1997; 97US-0056631P.
PR 22-AUG-1997; 97US-0056632P.
PR 22-AUG-1997; 97US-0056636P.
PR 22-AUG-1997; 97US-0056637P.
PR 22-AUG-1997; 97US-0056662P.
PR 22-AUG-1997; 97US-0056664P.
PR 22-AUG-1997; 97US-0056845P.
PR 22-AUG-1997; 97US-0056862P.
PR 22-AUG-1997; 97US-0056864P.
PR 22-AUG-1997; 97US-0056872P.
PR 22-AUG-1997; 97US-0056874P.
PR 22-AUG-1997; 97US-0056875P.
PR 22-AUG-1997; 97US-0056876P.
PR 22-AUG-1997; 97US-0056877P.
PR 22-AUG-1997; 97US-0056878P.
PR 22-AUG-1997; 97US-0056879P.
PR 22-AUG-1997; 97US-0056880P.
PR 22-AUG-1997; 97US-0056881P.
PR 22-AUG-1997; 97US-0056882P.
PR 22-AUG-1997; 97US-0056884P.
PR 22-AUG-1997; 97US-0056886P.

PR 22-AUG-1997; 97US-0056887P.
PR 22-AUG-1997; 97US-0056888P.
PR 22-AUG-1997; 97US-0056889P.
PR 22-AUG-1997; 97US-0056892P.
PR 22-AUG-1997; 97US-0056893P.
PR 22-AUG-1997; 97US-0056894P.
PR 22-AUG-1997; 97US-0056903P.
PR 22-AUG-1997; 97US-0056908P.
PR 22-AUG-1997; 97US-0056909P.
PR 22-AUG-1997; 97US-0056910P.
PR 22-AUG-1997; 97US-0056911P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057669P.
PR 05-SEP-1997; 97US-0057761P.
PR 12-SEP-1997; 97US-0058785P.
PR 09-OCT-1997; 97US-0061660P.
PR 06-MAR-1998; 98WO-US004493.
PR 08-SEP-1998; 98US-00149476.
PR 17-MAR-2000; 2000US-0190068P.
XX
PA (RUBE/) RUBEN S M.
PA (ROSE/) ROSEN C A.
PA (SOPP/) SOPPET D R.
PA (CART/) CARTER K C.
PA (BEDN/) BEDNARIK D P.
PA (ENDR/) ENDRESS G A.
PA (YUGG/) YU G.
PA (NIJJ/) NI J.
PA (FENG/) FENG P.
PA (YOUN/) YOUNG P E.
PA (GREE/) GREENE J M.
PA (FERR/) FERRIE A M.
PA (DUAN/) DUAN D R.
PA (HUJJ/) HU J.
PA (FLOR/) FLORENCE K A.
PA (OLSE/) OLSEN H S.
PA (FISC/) FISCHER C L.
PA (EBNE/) EBNER R.
PA (BREW/) BREWER L A.
PA (MOOR/) MOORE P A.
PA (SHIY/) SHI Y.
PA (LAFL/) LAFLEUR D W.
PA (LIYY/) LI Y.
PA (ZENG/) ZENG Z.
PA (KYAW/) KYAW H.
XX
PI Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP;
PI Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferrie AM;
PI Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R;
PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;
XX
DR WPI; 2003-521800/49.
DR N-PSDB; ACD82869.
XX
PT New genes and its encoded prostate cancer antigen proteins, useful for
PT preventing, treating, ameliorating or diagnosing e.g. prostate cancers,
PT thymic hypoplasia, multiple sclerosis, AIDS, angina pectoris or cerebral
PT ischemia.
XX
PS Claim 3; SEQ ID NO 548; 260pp; English.
XX
CC The present invention relates to the isolation of novel human secreted
CC proteins and the polynucleotide sequences encoding them. The invention
CC also discloses vectors, host cells, antibodies, and recombinant methods
CC for producing human secreted proteins. The polypeptide and polynucleotide
CC sequences for the secreted proteins are useful for preventing, treating,
CC ameliorating or diagnosing medical conditions such as hyperproliferative
CC disorders (e.g. leukaemia or breast cancers), wounds, reproductive
CC disorders, blood-related disorders (e.g. haemophilia or
CC thrombocytopaenia), immunodeficiencies (e.g. Wiskott-Aldrich syndrome or
CC thymic hypoplasia), autoimmune disorders (e.g. graft-versus-host disease,
CC multiple sclerosis or Hashimoto's thyroiditis), allergies (e.g. asthma),
CC viral or bacterial or fungal infections (e.g. AIDS or sepsis), renal
CC disorders (e.g. kidney failure), cardiovascular disorders (e.g. angina
CC pectoris, cerebral ischaemia or congenital heart defects), respiratory
CC disorders, neurological disorders (e.g. Alzheimer's disease or
CC Parkinson's disease), and inflammations (e.g. Crohn's disease). The
CC polynucleotide or polypeptide may also be used as vaccine adjuvants.
CC ABO34374-ABO34815 represent human secreted proteins or their fragments.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/psipsDIDEntry.html
XX
SQ Sequence 14 AA;

Query Match 36.4%; Score 28; DB 6; Length 14;
Best Local Similarity 36.4%; Pred. No. 8.5e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FEVVTFLKDVL 12
|: ::: ||:|
Db 2 FDFLSYFKDLL 12

RESULT 14
ADI23263
ID ADI23263 standard; protein; 14 AA.
XX
AC ADI23263;
XX
DT 22-APR-2004 (first entry)
XX
DE Novel human secreted protein seq id 548.
XX
KW cytostatic; gene therapy; cancer; human; secreted protein.
XX
OS Homo sapiens.
XX
PN US2003175858-A1.
XX
PD 18-SEP-2003.
XX
PF 18-JUN-2001; 2001US-00882171.
XX
PR 07-MAR-1997; 97US-0038621P.
PR 07-MAR-1997; 97US-0040162P.
PR 07-MAR-1997; 97US-0040163P.
PR 07-MAR-1997; 97US-0040333P.
PR 07-MAR-1997; 97US-0040334P.
PR 07-MAR-1997; 97US-0040336P.
PR 07-MAR-1997; 97US-0040626P.
PR 11-APR-1997; 97US-0043311P.
PR 11-APR-1997; 97US-0043312P.
PR 11-APR-1997; 97US-0043313P.
PR 11-APR-1997; 97US-0043314P.
PR 11-APR-1997; 97US-0043315P.
PR 11-APR-1997; 97US-0043568P.
PR 11-APR-1997; 97US-0043569P.
PR 11-APR-1997; 97US-0043576P.
PR 11-APR-1997; 97US-0043578P.
PR 11-APR-1997; 97US-0043580P.
PR 11-APR-1997; 97US-0043669P.
PR 11-APR-1997; 97US-0043670P.
PR 11-APR-1997; 97US-0043671P.
PR 11-APR-1997; 97US-0043672P.
PR 11-APR-1997; 97US-0043674P.
PR 23-MAY-1997; 97US-0047492P.
PR 23-MAY-1997; 97US-0047500P.
PR 23-MAY-1997; 97US-0047501P.
PR 23-MAY-1997; 97US-0047502P.
PR 23-MAY-1997; 97US-0047503P.
PR 23-MAY-1997; 97US-0047581P.
PR 23-MAY-1997; 97US-0047582P.
PR 23-MAY-1997; 97US-0047583P.
PR 23-MAY-1997; 97US-0047584P.
PR 23-MAY-1997; 97US-0047585P.

PR 23-MAY-1997; 97US-0047586P.
PR 23-MAY-1997; 97US-0047587P.
PR 23-MAY-1997; 97US-0047588P.
PR 23-MAY-1997; 97US-0047589P.
PR 23-MAY-1997; 97US-0047590P.
PR 23-MAY-1997; 97US-0047592P.
PR 23-MAY-1997; 97US-0047593P.
PR 23-MAY-1997; 97US-0047594P.
PR 23-MAY-1997; 97US-0047595P.
PR 23-MAY-1997; 97US-0047596P.
PR 23-MAY-1997; 97US-0047597P.
PR 23-MAY-1997; 97US-0047598P.
PR 23-MAY-1997; 97US-0047599P.
PR 23-MAY-1997; 97US-0047600P.
PR 23-MAY-1997; 97US-0047601P.
PR 23-MAY-1997; 97US-0047612P.
PR 23-MAY-1997; 97US-0047613P.
PR 23-MAY-1997; 97US-0047614P.
PR 23-MAY-1997; 97US-0047615P.
PR 23-MAY-1997; 97US-0047617P.
PR 23-MAY-1997; 97US-0047618P.
PR 23-MAY-1997; 97US-0047632P.
PR 23-MAY-1997; 97US-0047633P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048974P.
PR 13-JUN-1997; 97US-0049610P.
PR 08-JUL-1997; 97US-0051926P.
PR 16-JUL-1997; 97US-0052874P.
PR 18-AUG-1997; 97US-0055724P.
PR 22-AUG-1997; 97US-0056630P.
PR 22-AUG-1997; 97US-0056631P.
PR 22-AUG-1997; 97US-0056632P.
PR 22-AUG-1997; 97US-0056636P.
PR 22-AUG-1997; 97US-0056637P.
PR 22-AUG-1997; 97US-0056662P.
PR 22-AUG-1997; 97US-0056664P.
PR 22-AUG-1997; 97US-0056845P.
PR 22-AUG-1997; 97US-0056862P.
PR 22-AUG-1997; 97US-0056864P.
PR 22-AUG-1997; 97US-0056872P.
PR 22-AUG-1997; 97US-0056874P.
PR 22-AUG-1997; 97US-0056875P.
PR 22-AUG-1997; 97US-0056876P.
PR 22-AUG-1997; 97US-0056877P.
PR 22-AUG-1997; 97US-0056878P.
PR 22-AUG-1997; 97US-0056879P.
PR 22-AUG-1997; 97US-0056880P.
PR 22-AUG-1997; 97US-0056881P.
PR 22-AUG-1997; 97US-0056882P.
PR 22-AUG-1997; 97US-0056884P.
PR 22-AUG-1997; 97US-0056886P.
PR 22-AUG-1997; 97US-0056887P.
PR 22-AUG-1997; 97US-0056888P.
PR 22-AUG-1997; 97US-0056889P.
PR 22-AUG-1997; 97US-0056892P.
PR 22-AUG-1997; 97US-0056893P.
PR 22-AUG-1997; 97US-0056894P.
PR 22-AUG-1997; 97US-0056903P.
PR 22-AUG-1997; 97US-0056908P.
PR 22-AUG-1997; 97US-0056909P.
PR 22-AUG-1997; 97US-0056910P.
PR 22-AUG-1997; 97US-0056911P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057669P.
PR 05-SEP-1997; 97US-0057761P.
PR 12-SEP-1997; 97US-0058785P.
PR 09-OCT-1997; 97US-0061660P.
PR 06-MAR-1998; 98WO-US004493.
PR 08-SEP-1998; 98US-00149476.
PR 17-MAR-2000; 2000US-0190068P.
PR 16-MAR-2001; 2001US-00809391.
XX
PA (RUBE/) RUBEN S M.
PA (ROSE/) ROSEN C A.
PA (SOPP/) SOPPET D R.
PA (CART/) CARTER K C.
PA (BEDN/) BEDNARIK D P.
PA (ENDR/) ENDRESS G A.
PA (YUGG/) YU G.
PA (NIJJ/) NI J.
PA (FENG/) FENG P.
PA (YOUN/) YOUNG P E.
PA (GREE/) GREENE J M.
PA (FERR/) FERRIE A M.
PA (DUAN/) DUAN D R.
PA (HUJJ/) HU J.
PA (FLOR/) FLORENCE K A.
PA (OLSE/) OLSEN H S.
PA (FISC/) FISCHER C L.
PA (EBNE/) EBNER R.
PA (BREW/) BREWER L A.
PA (MOOR/) MOORE P A.
PA (SHIY/) SHI Y.
PA (LAFL/) LAFLEUR D W.
PA (LIYY/) LI Y.
PA (ZENG/) ZENG Z.
PA (KYAW/) KYAW H.
XX
PI Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP;
PI Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferrie AM;
PI Duan DR, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R;
PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;
XX
DR WPI; 2003-898535/82.
DR N-PSDB; ADI22954.
XX
PT New nucleic acid molecule, useful for preparing a medicament for
PT diagnosing, preventing, treating or ameliorating a medical condition
PT e.g., cancer.
XX
PS Claim 11; SEQ ID NO 548; 256pp; English.
XX
CC The invention describes an isolated nucleic acid comprising a sequence
CC having 95 % identity with: a polynucleotide fragment of a sequence not
CC given in the specification, or its allelic variant; a polynucleotide
CC fragment of the cDNA sequence; a polynucleotide sequence encoding a
CC polypeptide, or its fragment, domain, epitope or species homologue; or a
CC polynucleotide that hybridises under stringent conditions to any one of
CC the sequences of (a)-(c). The nucleic acid is useful for preparing a
CC medicament for diagnosing, preventing, treating or ameliorating a medical
CC condition e.g., cancer. The is the amino acid sequence of a novel human
CC secreted protein of the invention.
XX
SQ Sequence 14 AA;

Query Match 36.4%; Score 28; DB 7; Length 14;
Best Local Similarity 36.4%; Pred. No. 8.5e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FEVVTFLKDVL 12
|: ::: ||:|
Db 2 FDFLSYFKDLL 12

RESULT 15
ADH74265
ID ADH74265 standard; protein; 14 AA.
XX
AC ADH74265;
XX
DT 25-MAR-2004 (first entry)
XX
DE Human secreted protein #229.
XX
KW human; secreted protein; cancer; haematopoietic disorder;
KW endocrine disorder; immune system disease; inflammatory disorder.

XX
OS Homo sapiens.
XX
PN US2003225248-A1.
XX
PD 04-DEC-2003.
XX
PF 10-JUN-2002; 2002US-00164861.
XX
PR 07-MAR-1997; 97US-0038621P.
PR 07-MAR-1997; 97US-0040161P.
PR 07-MAR-1997; 97US-0040162P.
PR 07-MAR-1997; 97US-0040163P.
PR 07-MAR-1997; 97US-0040333P.
PR 07-MAR-1997; 97US-0040334P.
PR 07-MAR-1997; 97US-0040336P.
PR 07-MAR-1997; 97US-0040626P.
PR 11-APR-1997; 97US-0043311P.
PR 11-APR-1997; 97US-0043312P.
PR 11-APR-1997; 97US-0043313P.
PR 11-APR-1997; 97US-0043314P.
PR 11-APR-1997; 97US-0043315P.
PR 11-APR-1997; 97US-0043568P.
PR 11-APR-1997; 97US-0043569P.
PR 11-APR-1997; 97US-0043576P.
PR 11-APR-1997; 97US-0043578P.
PR 11-APR-1997; 97US-0043580P.
PR 11-APR-1997; 97US-0043669P.
PR 11-APR-1997; 97US-0043670P.
PR 11-APR-1997; 97US-0043671P.
PR 11-APR-1997; 97US-0043672P.
PR 11-APR-1997; 97US-0043674P.
PR 23-MAY-1997; 97US-0047492P.
PR 23-MAY-1997; 97US-0047500P.
PR 23-MAY-1997; 97US-0047501P.
PR 23-MAY-1997; 97US-0047502P.
PR 23-MAY-1997; 97US-0047503P.
PR 23-MAY-1997; 97US-0047581P.
PR 23-MAY-1997; 97US-0047582P.
PR 23-MAY-1997; 97US-0047583P.
PR 23-MAY-1997; 97US-0047584P.
PR 23-MAY-1997; 97US-0047585P.
PR 23-MAY-1997; 97US-0047586P.
PR 23-MAY-1997; 97US-0047587P.
PR 23-MAY-1997; 97US-0047588P.
PR 23-MAY-1997; 97US-0047589P.
PR 23-MAY-1997; 97US-0047590P.
PR 23-MAY-1997; 97US-0047592P.
PR 23-MAY-1997; 97US-0047593P.
PR 23-MAY-1997; 97US-0047594P.
PR 23-MAY-1997; 97US-0047595P.
PR 23-MAY-1997; 97US-0047596P.
PR 23-MAY-1997; 97US-0047597P.
PR 23-MAY-1997; 97US-0047598P.
PR 23-MAY-1997; 97US-0047599P.
PR 23-MAY-1997; 97US-0047600P.
PR 23-MAY-1997; 97US-0047601P.
PR 23-MAY-1997; 97US-0047612P.
PR 23-MAY-1997; 97US-0047613P.
PR 23-MAY-1997; 97US-0047614P.
PR 23-MAY-1997; 97US-0047615P.
PR 23-MAY-1997; 97US-0047617P.
PR 23-MAY-1997; 97US-0047618P.
PR 23-MAY-1997; 97US-0047632P.
PR 23-MAY-1997; 97US-0047633P.
PR 06-JUN-1997; 97US-0048964P.
PR 06-JUN-1997; 97US-0048974P.
PR 13-JUN-1997; 97US-0049610P.
PR 08-JUL-1997; 97US-0051926P.
PR 16-JUL-1997; 97US-0052874P.
PR 18-AUG-1997; 97US-0055724P.
PR 22-AUG-1997; 97US-0056630P.
PR 22-AUG-1997; 97US-0056631P.
PR 22-AUG-1997; 97US-0056632P.
PR 22-AUG-1997; 97US-0056636P.
PR 22-AUG-1997; 97US-0056637P.
PR 22-AUG-1997; 97US-0056662P.
PR 22-AUG-1997; 97US-0056664P.
PR 22-AUG-1997; 97US-0056845P.
PR 22-AUG-1997; 97US-0056862P.
PR 22-AUG-1997; 97US-0056864P.
PR 22-AUG-1997; 97US-0056872P.
PR 22-AUG-1997; 97US-0056874P.
PR 22-AUG-1997; 97US-0056875P.
PR 22-AUG-1997; 97US-0056876P.
PR 22-AUG-1997; 97US-0056877P.
PR 22-AUG-1997; 97US-0056878P.
PR 22-AUG-1997; 97US-0056879P.
PR 22-AUG-1997; 97US-0056880P.
PR 22-AUG-1997; 97US-0056881P.
PR 22-AUG-1997; 97US-0056882P.
PR 22-AUG-1997; 97US-0056884P.
PR 22-AUG-1997; 97US-0056886P.
PR 22-AUG-1997; 97US-0056887P.
PR 22-AUG-1997; 97US-0056888P.
PR 22-AUG-1997; 97US-0056889P.
PR 22-AUG-1997; 97US-0056892P.
PR 22-AUG-1997; 97US-0056893P.
PR 22-AUG-1997; 97US-0056894P.
PR 22-AUG-1997; 97US-0056903P.
PR 22-AUG-1997; 97US-0056908P.
PR 22-AUG-1997; 97US-0056909P.
PR 22-AUG-1997; 97US-0056910P.
PR 22-AUG-1997; 97US-0056911P.
PR 05-SEP-1997; 97US-0057650P.
PR 05-SEP-1997; 97US-0057669P.
PR 05-SEP-1997; 97US-0057761P.
PR 12-SEP-1997; 97US-0058785P.
PR 02-OCT-1997; 97US-0061060P.
PR 06-MAR-1998; 98WO-US004493.
PR 08-SEP-1998; 98US-00149476.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Rosen CA, Soppet DR, Carter KC, Bednarik DP;
PI Endress GA, Yu G, Ni J, Feng P, Young PE, Greene JM, Ferrie AM;
PI Duan R, Hu J, Florence KA, Olsen HS, Fischer CL, Ebner R;
PI Brewer LA, Moore PA, Shi Y, Lafleur DW, Li Y, Zeng Z, Kyaw H;
XX
DR WPI; 2004-131264/13.
DR N-PSDB; ADH73956.
XX
PT Isolated nucleic acid molecules encoding human secreted proteins, useful
PT for preventing, diagnosing and treating disorders associated with
PT aberrant expression and activity.
XX
PS Claim 11; SEQ ID NO 548; 142pp; English.
XX
CC The invention relates to isolated nucleic acid molecules and the human
CC secreted proteins (SPs) they encode. The proteins and nucleic acids may
CC be used in the prevention, diagnosis and treatment of diseases associated
CC with inappropriate SP expression e.g. cancer, haematopoietic disorders,
CC endocrine disorders, diseases of the immune system, inflammatory
CC disorders and many others. Full details of disorders that may be
CC prevented, diagnosed and/or treated by the above methods are given in the
CC specification. The nucleic acid molecules may be used to produce their
CC proteins. The nucleic acid and it's complementary sequences may also be
CC used as DNA probes in diagnostic assays to detect and quantitate the
CC presence of similar nucleic acids in samples, and therefore which
CC patients may be in need of restorative therapy. The SPs may also be used
CC as antigens in the production of antibodies against the proteins and in
CC assays to identify modulators of SP expression and activity. The anti-SP
CC antibodies and antagonists may also be used to down regulate expression
CC and activity. The anti-SP antibodies may also be used as diagnostic
CC agents for detecting the presence of the proteins in samples (e.g. by
CC enzyme linked immunosorbant assay (ELISA)). The present sequence

CC represents the amino acid sequence of a human secreted protein.
XX
SQ Sequence 14 AA;

Query Match 36.4%; Score 28; DB 8; Length 14;
Best Local Similarity 36.4%; Pred. No. 8.5e+02;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 FEVVTFLKDVL 12
              |: ::: ||:|
Db          2 FDFLSYFKDLL 12
```

Search completed: February 22, 2005, 09:24:46
Job time : 67.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-10
Perfect score: 72
Sequence: 1 DRLLMLFAKDVVSRN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	23	31.9	13	2	PN0168	phosphopyruvate hy
2	23	31.9	14	2	A61308	hemocyanin chain 2
3	23	31.9	14	2	D61308	hemocyanin chain 5
4	22	30.6	13	2	S29488	GTP-binding protei
5	22	30.6	15	2	PA0093	enniatin synthetas
6	21	29.2	11	2	S42449	ant1 protein - pha
7	20	27.8	14	2	E61308	hemocyanin chain 3
8	20	27.8	14	2	PH1356	Ig heavy chain DJ
9	20	27.8	14	2	G61308	hemocyanin chain 3
10	19	26.4	12	2	S01222	translation elonga
11	19	26.4	15	2	PH1590	Ig H chain V-D-J r
12	18	25.0	11	2	B49164	chromogranin-B - r
13	18	25.0	11	2	PC2372	58K heat shock pro
14	18	25.0	14	2	PS0371	hypothetical prote
15	18	25.0	14	2	S29789	hypothetical prote
16	18	25.0	14	2	F83754	hypothetical prote
17	18	25.0	15	2	PA0052	protein QF200015 -
18	18	25.0	15	2	D28587	T-cell receptor be
19	18	25.0	15	2	PA0075	fructose-bisphosph
20	17	23.6	11	2	PT0229	Ig heavy chain CDR
21	17	23.6	12	4	PC2121	aminotransferase c
22	17	23.6	14	1	QMVHP2	mastoparan C - Eur
23	17	23.6	14	2	S19803	ubiquitin - potato
24	17	23.6	15	2	S61438	hypothetical prote
25	17	23.6	15	2	PA0102	fructose-bisphosph
26	17	23.6	15	2	PA0063	ubiquitin - fungus
27	17	23.6	15	2	I46909	voltage-dependent
28	16	22.2	6	2	I59142	platelet-derived g
29	16	22.2	8	2	E60588	sperm-activating p
30	16	22.2	11	2	S33782	acetolactate synth
31	16	22.2	12	2	S47360	T-cell antigen rec
32	16	22.2	12	2	PH1587	Ig H chain V-D-J r
33	16	22.2	12	2	B61497	seed protein ws-17
34	16	22.2	12	2	A34858	proteinase E - bla
35	16	22.2	13	2	PN0122	O1L protein - vacc
36	16	22.2	13	2	A61288	spore proteinase g
37	16	22.2	13	2	S36887	ribosomal protein
38	16	22.2	13	2	S01043	glutamate-ammonia
39	16	22.2	14	2	E33098	214K exoantigen (v
40	16	22.2	14	2	PH1327	Ig heavy chain DJ
41	16	22.2	14	2	JS0272	hypothetical 1.5K
42	16	22.2	14	2	PH1617	Ig H chain V-D-J r
43	16	22.2	14	2	S14336	mastoparan B - hor
44	16	22.2	15	2	A60834	angiotensin I prec
45	16	22.2	15	2	S55312	TSH protein beta c

ALIGNMENTS

RESULT 1
PN0168
phosphopyruvate hydratase (EC 4.2.1.11) - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 09-Jul-2004
C;Accession: PN0168
R;Fukaya, N.; Chow, L.P.; Sugiura, Y.; Tsugita, A.; Ueno, Y.; Tabuchi, K.
submitted to JIPID, May 1994
A;Description: Two dimensional polyacrylamide gel electrophoresis of Fusarium sporotrici
A;Reference number: PN0160
A;Accession: PN0168
A;Molecule type: protein
A;Residues: 1-13 <FUK>
A;Cross-references: UNIPROT:Q7M4Y6
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 31.9%; Score 23; DB 2; Length 13;
Best Local Similarity 25.0%; Pred. No. 6e+02;
Matches 3; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

```
Qy          3 LLMLFAKDVVSR 14
              ::::||: | :
Db          2 IVLVFARSVYDK 13
```

RESULT 2
A61308
hemocyanin chain 2 - Sahara scorpion (fragment)
C;Species: Androctonus australis (Sahara scorpion)
C;Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 09-Jul-2004
C;Accession: A61308
R;Jolles, J.; Jolles, P.; Lamy, J.; Lamy, J.
FEBS Lett. 106, 289-291, 1979
A;Title: Structural characterization of seven different subunits in Androctonus australi
A;Reference number: A61308; MUID:80047238; PMID:499512
A;Accession: A61308
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <JOL>
A;Cross-references: UNIPROT:Q7M488

Query Match 31.9%; Score 23; DB 2; Length 14;
Best Local Similarity 57.1%; Pred. No. 6.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 DRLLMLF 7
              ||:: ||
Db          7 DRIIPLF 13
```

RESULT 3
D61308

```
hemocyanin chain 5A - Sahara scorpion (fragment)
C;Species: Androctonus australis (Sahara scorpion)
C;Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 09-Jul-2004
C;Accession: D61308
R;Jolles, J.; Jolles, P.; Lamy, J.; Lamy, J.
FEBS Lett. 106, 289-291, 1979
A;Title: Structural characterization of seven different subunits in Androctonus australi
A;Reference number: A61308; MUID:80047238; PMID:499512
A;Accession: D61308
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <JOL>
A;Cross-references: UNIPROT:Q7M486

  Query Match             31.9%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 6.5e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DRLLMLF 7
              :||| ||
Db          7 ERLLPLF 13


RESULT 4
S29488
GTP-binding protein o-rab3 - marbled electric ray (fragment)
C;Species: Torpedo marmorata (marbled electric ray)
C;Date: 22-Nov-1993 #sequence_revision 27-Feb-1997 #text_change 13-Mar-1997
C;Accession: S29488
R;Volknandt, W.; Pevsner, J.; Elferink, L.A.; Scheller, R.H.
FEBS Lett. 317, 53-56, 1993
A;Title: Association of three small GTP-binding proteins with cholinergic synaptic vesic
A;Reference number: S29485; MUID:93154521; PMID:8428634
A;Accession: S29488
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-13 <VOL>

  Query Match             30.6%;  Score 22;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 9.3e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 KDVVSRN 15
              || | :|
Db          1 KDAVDQN 7


RESULT 5
PA0093
enniatin synthetase - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0093
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0093
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Z6

  Query Match             30.6%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 LFAKD 10
              |:|||
Db          8 LYAKD 12


RESULT 6
S42449
ant1 protein - phage P7
C;Species: phage P7
C;Date: 07-Sep-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: S42449
R;Citron, M.; Schuster, H.
Cell 62, 591-598, 1990
A;Title: The c4 repressors of bacteriophages P1 and P7 are antisense RNAs.
A;Reference number: S42448; MUID:90335968; PMID:1696181
A;Accession: S42449
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-11 <CIT>
A;Cross-references: UNIPROT:Q38415; EMBL:M35139; NID:g215705; PIDN:AAA32437.1; PID:g2157

  Query Match             29.2%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KDVVSRN 15
              | :|:||
Db          3 KPLVTRN 9


RESULT 7
E61308
hemocyanin chain 3B - Sahara scorpion (fragment)
C;Species: Androctonus australis (Sahara scorpion)
C;Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
C;Accession: E61308
R;Jolles, J.; Jolles, P.; Lamy, J.; Lamy, J.
FEBS Lett. 106, 289-291, 1979
A;Title: Structural characterization of seven different subunits in Androctonus australi
A;Reference number: A61308; MUID:80047238; PMID:499512
A;Accession: E61308
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <JOL>

  Query Match             27.8%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLLMLF 7
              |:| ||
Db          8 RILKLF 13


RESULT 8
PH1356
Ig heavy chain DJ region (clone C178-97) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1356
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A;Reference number: PH1302; MUID:93094761; PMID:1460419
A;Accession: PH1356
A;Molecule type: DNA
A;Residues: 1-14 <WAS>
C;Keywords: heterotetramer; immunoglobulin

  Query Match             27.8%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   33.3%;  Pred. No. 2.4e+03;
  Matches    3;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 LLMLFAKDV 11
              : |::  ||
Db          1 ITMIYGMDV 9


RESULT 9
G61308
```

```
hemocyanin chain 3C - Sahara scorpion (fragment)
C;Species: Androctonus australis (Sahara scorpion)
C;Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 09-Jul-2004
C;Accession: G61308
R;Jolles, J.; Jolles, P.; Lamy, J.; Lamy, J.
FEBS Lett. 106, 289-291, 1979
A;Title: Structural characterization of seven different subunits in Androctonus australi
A;Reference number: A61308; MUID:80047238; PMID:499512
A;Accession: G61308
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <JOL>
A;Cross-references: UNIPROT:Q7M487

  Query Match             27.8%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLLMLF 7
              |:| ||
Db          8 RILSLF 13


RESULT 10
S01222
translation elongation factor EF-Tu - Pseudomonas aeruginosa (fragment)
C;Species: Pseudomonas aeruginosa
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C;Accession: S01222
R;Hughes, M.A.; Jones, D.S.
Nucleic Acids Res. 16, 7193, 1988
A;Title: A fragment of the Pseudomonas aeruginosa genome contains five tRNA genes, four
A;Reference number: S01222; MUID:88303352; PMID:3136442
A;Accession: S01222
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-12 <HUG>
A;Cross-references: UNIPROT:P09591; EMBL:X07950; NID:g45426; PIDN:CAA30775.1; PID:g45427
C;Genetics:
A;Gene: tufB
C;Superfamily: translation elongation factor Tu; translation elongation factor Tu homolo
C;Keywords: GTP binding; protein biosynthesis

  Query Match             26.4%;  Score 19;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 3.1e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          8 AKDVVSRN 15
              ||:   ||
Db          2 AKEKFERN 9


RESULT 11
PH1590
Ig H chain V-D-J region (wild-type clone 141) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1590
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609; PMID:8315387
A;Accession: PH1590
A;Molecule type: DNA
A;Residues: 1-15 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin

  Query Match             26.4%;  Score 19;  DB 2;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 3.9e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          8 AKDVVSR 14
              |:|:: |
Db          2 ARDMILR 8


RESULT 12
B49164
chromogranin-B - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 31-Oct-1997
C;Accession: B49164
R;Nielsen, E.; Welinder, B.S.; Madsen, O.D.
Endocrinology 129, 3147-3156, 1991
A;Title: Chromogranin-B, a putative precursor of eight novel rat glucagonoma peptides th
A;Reference number: A49164; MUID:92063871; PMID:1954895
A;Accession: B49164
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <NIE>
A;Note: sequence extracted from NCBI backbone (NCBIP:66370)
C;Superfamily: chromogranin B precursor

  Query Match             25.0%;  Score 18;  DB 2;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 4.4e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 FAKDV 11
              |::||
Db          2 FSEDV 6


RESULT 13
PC2372
58K heat shock protein groEL [similarity] - Bacillus cereus (strain ts-4) (fragment)
C;Species: Bacillus cereus
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: PC2372
R;Matsuno, K.; Miyamoto, T.; Yamaguchi, K.; Sayed, M.A.; Kajiwara, T.; Hatano, S.
Biosci. Biotechnol. Biochem. 59, 231-235, 1995
A;Title: Identification of DNA-binding proteins changed after induction of sporulation
A;Reference number: PC2369; MUID:95218265; PMID:7766022
A;Accession: PC2372
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <MAS>
C;Keywords: heat shock; molecular chaperone; stress-induced protein

  Query Match             25.0%;  Score 18;  DB 2;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 4.4e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 AKDV 11
              |||:
Db          1 AKDI 4


RESULT 14
PS0371
hypothetical protein (psaC region) - Synechococcus sp. (fragment)
C;Species: Synechococcus sp.
C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 08-Oct-1999
C;Accession: PS0371
R;Rhiel, E.; Stirewalt, V.L.; Gasparich, G.E.; Bryant, D.A.
Gene 112, 123-128, 1992
A;Title: The psaC genes of Synechococcus sp. PCC7002 and Cyanophora paradoxa: cloning a
A;Reference number: JS0694; MUID:92201692; PMID:1551590
A;Accession: PS0371
A;Molecule type: DNA
A;Residues: 1-14 <RHI>
A;Cross-references: GB:M86238; NID:g154574; PIDN:AAA27351.1; PID:g552030

  Query Match             25.0%;  Score 18;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 5.7e+03;
  Matches    4;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy          7 FAKDVVSR 14
              |  ||  |
Db          4 FKLDVTGR 11
```

RESULT 15
S29789
hypothetical protein - Thermoplasma acidophilum (fragment)
C;Species: Thermoplasma acidophilum
C;Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 26-May-1995
C;Accession: S29789
R;Bright, J.R.; Byrom, D.; Danson, M.J.; Hough, D.W.; Towner, P.
Eur. J. Biochem. 211, 549-554, 1993
A;Title: Cloning, sequencing and expression of the gene encoding glucose dehydrogenase f
A;Reference number: S29788; MUID:93170285; PMID:8436115
A;Accession: S29789
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-14 <BRI>
A;Cross-references: EMBL:X59788

```
  Query Match             25.0%;  Score 18;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 5.7e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KDVVSR 14
              ||:: |
Db          2 KDLLKR 7
```

Search completed: February 22, 2005, 09:46:28
Job time : 12.1333 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-10
Perfect score: 72
Sequence: 1 DRLLMLFAKDVVSRN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	36.1	12	1	UN39_CLOPA	P81359 clostridium
2	25	34.7	15	2	Q9S8D3	Q9s8d3 cynara card
3	23	31.9	13	2	Q7M4Y6	Q7m4y6 fusarium sp
4	23	31.9	14	2	Q7M486	Q7m486 androctonus
5	23	31.9	14	2	Q7M488	Q7m488 androctonus
6	23	31.9	15	1	GTE2_PSEUO	P83000 pseudomonas
7	22	30.6	12	2	Q9UR28	Q9ur28 cryptococcu
8	22	30.6	14	2	Q15998	Q15998 homo sapien
9	22	30.6	15	2	Q7M4Z6	Q7m4z6 fusarium sp
10	21	29.2	11	2	Q38415	Q38415 bacteriopha
11	21	29.2	11	2	Q9R7U8	Q9r7u8 pseudomonas
12	21	29.2	13	2	Q6TUI7	Q6tui7 ascaris suu
13	21	29.2	15	2	Q950P6	Q950p6 rhizophydiu
14	20	27.8	10	1	GS09_BACSU	P80243 bacillus su
15	20	27.8	14	2	Q7M487	Q7m487 androctonus
16	20	27.8	14	2	Q7RR89	Q7rr89 plasmodium
17	20	27.8	14	2	Q661E1	Q661e1 borrelia ga
18	20	27.8	15	2	Q9UWH4	Q9uwh4 pyrococcus.
19	20	27.8	15	2	Q9UWH6	Q9uwh6 thermococcu
20	19	26.4	9	2	Q7S182	Q7s182 neurospora
21	19	26.4	11	2	Q7RBB2	Q7rbb2 plasmodium
22	19	26.4	12	2	Q6JDK3	Q6jdk3 canis famil
23	19	26.4	12	2	Q9QZD0	Q9qzd0 mus musculu
24	19	26.4	13	1	CH60_CANFA	P49818 canis famil
25	19	26.4	13	2	Q8IZR0	Q8izr0 homo sapien
26	19	26.4	13	2	Q6LCB1	Q6lcb1 rattus norv
27	19	26.4	14	2	Q8J1G3	Q8j1g3 ashbya goss
28	19	26.4	14	2	Q8WWR7	Q8wwr7 homo sapien
29	19	26.4	14	2	Q9UHM5	Q9uhm5 homo sapien
30	19	26.4	15	2	Q8K1P5	Q8k1p5 sciurus vul
31	19	26.4	15	2	Q8K1W5	Q8k1w5 castor cana
32	19	26.4	15	2	Q9QXZ5	Q9qxz5 mus musculu
33	18.5	25.7	14	1	UN46_CLOPA	P81362 clostridium
34	18.5	25.7	15	2	Q9TNP0	Q9tnp0 mus sp. . m
35	18	25.0	10	1	MALE_KLEPN	Q05564 klebsiella
36	18	25.0	10	1	RCA_PINPS	P81084 pinus pinas
37	18	25.0	10	2	Q6LEM6	Q6lem6 homo sapien
38	18	25.0	10	2	O90344	O90344 gb virus c/
39	18	25.0	11	2	Q9QVH3	Q9qvh3 rattus sp.
40	18	25.0	11	2	Q66200	Q66200 transmissib
41	18	25.0	12	2	Q6EGT2	Q6egt2 stemphylium
42	18	25.0	12	2	Q6EGV5	Q6egv5 pleospora g
43	18	25.0	12	2	Q7KYL7	Q7kyl7 homo sapien
44	18	25.0	12	2	Q90QD1	Q90qd1 human immun
45	18	25.0	13	1	BP37_LEUMA	P81754 leucophaea

ALIGNMENTS

RESULT 1
UN39_CLOPA
ID UN39_CLOPA STANDARD; PRT; 12 AA.
AC P81359;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Unknown protein CP 39 from 2D-PAGE (Fragment).
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1501;
RN [1]
RP SEQUENCE.
RC STRAIN=W5;
RX MEDLINE=98291870; PubMed=9629918;
RA Flengsrud R., Skjeldal L.;
RT "Two-dimensional gel electrophoresis separation and N-terminal
RT sequence analysis of proteins from Clostridium pasteurianum W5.";
RL Electrophoresis 19:802-806(1998).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 5.4, its MW is: 29.5 kDa.
KW Direct protein sequencing.
FT NON_TER 12 12
SQ SEQUENCE 12 AA; 1432 MW; 940561E66BD2CB01 CRC64;

Query Match 36.1%; Score 26; DB 1; Length 12;
Best Local Similarity 36.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy          5 MLFAKDVVSRN 15
              |::: :||: |
Db          1 MIYSTEVVNMN 11
```

RESULT 2
Q9S8D3
ID Q9S8D3 PRELIMINARY; PRT; 15 AA.
AC Q9S8D3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CARDOSIN B (Fragment).
OS Cynara cardunculus (Cardoon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Carduoideae; Cardueae; Cynara.
OX NCBI_TaxID=4265;
RN [1]
RP SEQUENCE.
RX MEDLINE=96073661; PubMed=8540346;
RA Faro C., Verissimo P., Lin Y., Tang J., Pires E.;
RT "Cardosin A and B, aspartic proteases from the flowers of cardoon.";
RL Adv. Exp. Med. Biol. 362:373-377(1995).

```
SQ   SEQUENCE   15 AA;  1571 MW;  1DD25E8CB119379B CRC64;

  Query Match             34.7%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+03;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          8 AKDVVSRN 15
              |:|:|: |
Db          2 AEDIVNNN 9


RESULT 3
Q7M4Y6
ID   Q7M4Y6      PRELIMINARY;      PRT;    13 AA.
AC   Q7M4Y6;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Phosphopyruvate hydratase (EC 4.2.1.11) (Fragment).
OS   Fusarium sporotrichioides.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
OX   NCBI_TaxID=5514;
RN   [1]
RP   SEQUENCE.
RA   Fukaya N., Chow L.P., Sugiura Y., Tsugita A., Ueno Y., Tabuchi K.;
RL   Submitted (MAY-1994) to the PIR data bank.
DR   PIR; PN0168; PN0168.
DR   GO; GO:0004634; F:phosphopyruvate hydratase activity; IEA.
FT   NON_TER       1      1
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1481 MW;  CFEA8596A8A10DD9 CRC64;

  Query Match             31.9%;  Score 23;  DB 2;  Length 13;
  Best Local Similarity   25.0%;  Pred. No. 5.2e+03;
  Matches    3;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 LLMLFAKDVVSR 14
              ::::||: |  :
Db          2 IVLVFARSVYDK 13


RESULT 4
Q7M486
ID   Q7M486      PRELIMINARY;      PRT;    14 AA.
AC   Q7M486;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hemocyanin chain 5A (Fragment).
OS   Androctonus australis (Sahara scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthida; Buthoidea; Buthidae; Androctonus.
OX   NCBI_TaxID=6858;
RN   [1]
RP   SEQUENCE.
RA   Jolles J., Jolles P., Lamy J., Lamy J.;
RT   "Structural characterization of seven different subunits in
RT   Androctonus australis haemocyanin.";
RL   FEBS Lett. 106:289-291(1979).
DR   PIR; D61308; D61308.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1658 MW;  6FA4342770FEAF0B CRC64;

  Query Match             31.9%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 5.6e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DRLLMLF 7
              :||| ||
Db          7 ERLLPLF 13


RESULT 5
Q7M488
ID   Q7M488      PRELIMINARY;      PRT;    14 AA.
AC   Q7M488;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hemocyanin chain 2 (Fragment).
OS   Androctonus australis (Sahara scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthida; Buthoidea; Buthidae; Androctonus.
OX   NCBI_TaxID=6858;
RN   [1]
RP   SEQUENCE.
RA   Jolles J., Jolles P., Lamy J., Lamy J.;
RT   "Structural characterization of seven different subunits in
RT   Androctonus australis haemocyanin.";
RL   FEBS Lett. 106:289-291(1979).
DR   PIR; A61308; A61308.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1716 MW;  9B14357E859E400A CRC64;

  Query Match             31.9%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   57.1%;  Pred. No. 5.6e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 DRLLMLF 7
              ||:: ||
Db          7 DRIIPLF 13


RESULT 6
GTE2_PSEUO
ID   GTE2_PSEUO     STANDARD;      PRT;    15 AA.
AC   P83000;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Glutathione S-transferase (EC 2.5.1.18) (Fragment).
OS   Pseudomonas sp. (strain M1).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=95619;
RN   [1]
RP   SEQUENCE, FUNCTION, CATALYTIC ACTIVITY, SUBUNIT, AND SUBCELLULAR
RP   LOCATION.
RX   MEDLINE=21896940; PubMed=11900268; DOI=10.1016/S0923-2508(01)01293-1;
RA   Santos P.M., Mignogna G., Heipieper H.J., Zennaro E.;
RT   "Occurence and properties of glutathione S-transferases in phenol-
RT   degrading Pseudomonas strains.";
RL   Res. Microbiol. 153:89-98(2002).
CC   -!- FUNCTION: Conjugation of reduced glutathione to a wide number of
CC       exogenous and endogenous hydrophobic electrophiles.
CC   -!- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC   -!- SUBUNIT: Monomer and homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the GST superfamily.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0004364; F:glutathione transferase activity; NAS.
DR   GO; GO:0008152; P:metabolism; IC.
KW   Direct protein sequencing; Transferase.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1608 MW;  571C90DA7595077A CRC64;

  Query Match             31.9%;  Score 23;  DB 1;  Length 15;
  Best Local Similarity   38.5%;  Pred. No. 6e+03;
  Matches    5;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 LLMLFAKDVVSRN 15
              ||:: :|:: | |
Db          2 LLVIGSKNLSSTN 14
```

```
RESULT 7
Q9UR28
ID   Q9UR28      PRELIMINARY;      PRT;    12 AA.
AC   Q9UR28;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   STE12 alpha (Fragment).
GN   Name=STE12alpha;
OS   Cryptococcus bacillisporus (Filobasidiella neoformans var.
OS   bacillispora).
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC   Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX   NCBI_TaxID=37769;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99380307; PubMed=10449476;
RA   Halliday C.L., Bui T., Krockenberger M., Malik R., Ellis D.H.,
RA   Carter D.A.;
RT   "Presence of alpha and a mating types in environmental and clinical
RT   collections of Cryptococcus neoformans var. gattii strains from
RT   Australia.";
RL   J. Clin. Microbiol. 37:2920-2926(1999).
DR   EMBL; AF155346; AAF20371.1; -.
DR   EMBL; AF155343; AAF20368.1; -.
DR   EMBL; AF155344; AAF20369.1; -.
DR   EMBL; AF155345; AAF20370.1; -.
DR   EMBL; AF155342; AAF20367.1; -.
FT   NON_TER       1      1
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1477 MW;  460BA17AE8A729C7 CRC64;

  Query Match           30.6%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   57.1%;  Pred. No. 7.5e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 DRLLMLF 7
              ||| : |
Db          3 DRLFLTF 9


RESULT 8
Q15998
ID   Q15998      PRELIMINARY;      PRT;    14 AA.
AC   Q15998;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Neurofibromatosis type 1 gene protein (Fragment).
GN   Name=neurofibromatosis type 1 gene;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93216281; PubMed=8385067;
RA   Ainsworth P.J., Rodenhiser D.I., Costa M.T.;
RT   "Identification and characterization of sporadic and inherited
RT   mutations in exon 31 of the neurofibromatosis (NF1) gene.";
RL   Hum. Genet. 91:151-156(1993).
DR   EMBL; S57966; AAD13890.1; -.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1660 MW;  DC7FB5EB30767332 CRC64;

  Query Match           30.6%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 8.6e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 RLLMLF 7
              |:|:||
Db          6 RVLLLF 11


RESULT 9
Q7M4Z6
ID   Q7M4Z6      PRELIMINARY;      PRT;    15 AA.
AC   Q7M4Z6;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Enniatin synthetase (Fragment).
OS   Fusarium sporotrichioides.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
OX   NCBI_TaxID=5514;
RN   [1]
RP   SEQUENCE.
RA   Chow L.P., Fukaya N., Sugiura Y., Ueno Y., Tabuchi K., Tsugita A.;
RL   Submitted (OCT-1994) to the PIR data bank.
DR   PIR; PA0093; PA0093.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1828 MW;  5C31D6723D77F4D9 CRC64;

  Query Match           30.6%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 9.2e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 LFAKD 10
              |:|||
Db          8 LYAKD 12


RESULT 10
Q38415
ID   Q38415      PRELIMINARY;      PRT;    11 AA.
AC   Q38415;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ant1 protein (Fragment).
OS   Bacteriophage P7.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC   P1-like viruses.
OX   NCBI_TaxID=10682;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90335968; PubMed=1696181; DOI=10.1016/0092-8674(90)90023-8;
RA   Citron M., Schuster H.;
RT   "The c4 repressors of bacteriophages P1 and P7 are antisense RNAs.";
RL   Cell 62:591-598(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92319637; PubMed=1620606;
RA   Citron M., Schuster H.;
RT   "The c4 repressor of bacteriophage P1 is a processed 77 base antisense
RT   RNA.";
RL   Nucleic Acids Res. 20:3085-3090(1992).
DR   EMBL; M35139; AAA32437.1; -.
DR   PIR; S42449; S42449.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1315 MW;  38A55C6D11B2C737 CRC64;

  Query Match           29.2%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KDVVSRN 15
              | :|:||
Db          3 KPLVTRN 9


RESULT 11
```

```
Q9R7U8
ID   Q9R7U8      PRELIMINARY;      PRT;    11 AA.
AC   Q9R7U8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   DNR protein (a regulatory protein for the expression of the nir and
DE   nor genes), complete cds. (Fragment).
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=95226457; PubMed=7711073; DOI=10.1016/0167-4781(95)00018-C;
RA   Arai H., Igarashi Y., Kodama T.;
RT   "The structural genes for nitric oxide reductase from Pseudomonas
RT   aeruginosa.";
RL   Biochim. Biophys. Acta 1261:279-284(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAO1;
RX   MEDLINE=95394152; PubMed=7664887; DOI=10.1016/0014-5793(95)00885-D;
RA   Arai H., Igarashi Y., Kodama T.;
RT   "Expression of the nir and nor genes for denitrification of
RT   Pseudomonas aeruginosa requires a novel CRP/FNR-related
RT   transcriptional regulator, DNR, in addition to ANR.";
RL   FEBS Lett. 371:73-76(1995).
DR   EMBL; D50019; BAA08746.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   11 AA;  1543 MW;  DF363CAE141B5736 CRC64;

  Query Match             29.2%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+04;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 RLLMLFAK 9
              ||| |: :
Db          1 RLLQLYRR 8


RESULT 12
Q6TUI7
ID   Q6TUI7      PRELIMINARY;      PRT;    13 AA.
AC   Q6TUI7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   AF17 D (Fragment).
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID=6253;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nanda J.C., Stretton A.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY386839; AAQ90312.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   13 AA;  1531 MW;  18DA23119D6C79C4 CRC64;

  Query Match             29.2%;  Score 21;  DB 2;  Length 13;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+04;
  Matches    5;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DRLLMLFAK 9
              ||  | | |
Db          1 DRNFMNFGK 9


RESULT 13
Q950P6
ID   Q950P6      PRELIMINARY;      PRT;    15 AA.
AC   Q950P6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Orf15 (Fragment).
GN   Name=orf15;
OS   Rhizophydium sp. 136.
OG   Mitochondrion.
OC   Eukaryota; Fungi; Chytridiomycota; Chytridiales; Chytridiaceae;
OC   Rhizophydium.
OX   NCBI_TaxID=60187;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=136;
RX   MEDLINE=21851207; PubMed=11861890;
RA   Forget L., Ustinova J., Wang Z., Huss V.A.R., Lang B.F.;
RT   "Hyaloraphidium curvatum: a linear mitochondrial genome, tRNA editing,
RT   and an evolutionary link to lower fungi.";
RL   Mol. Biol. Evol. 19:310-319(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=136;
RA   Lang F.B.F.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF404306; AAK84262.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1721 MW;  99ABBEFDA9D6BCAE CRC64;

  Query Match             29.2%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+04;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 LLMLFA 8
              |::|||
Db          9 LMILFA 14


RESULT 14
GS09_BACSU
ID   GS09_BACSU     STANDARD;      PRT;    10 AA.
AC   P80243;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   General stress protein 9 (GSP9) (Fragment).
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1423;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=168 / IS58;
RX   MEDLINE=94282319; PubMed=8012595;
RA   Voelker U., Engelmann S., Maul B., Riethdorf S., Voelker A.,
RA   Schmid R., Mach H., Hecker M.;
RT   "Analysis of the induction of general stress proteins of Bacillus
RT   subtilis.";
RL   Microbiology 140:741-752(1994).
CC   -!- INDUCTION: By heat shock, salt stress, oxidative stress, glucose
CC       limitation and oxygen limitation.
CC   -!- CAUTION: Could not be found in the genome of B.subtilis 168.
KW   Direct protein sequencing; Heat shock.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1168 MW;  99766442D5A2C05A CRC64;

  Query Match             27.8%;  Score 20;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+04;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 AKDVVS 13
              ::|:||
```

```
Db          1 SRDIVS 6


RESULT 15
Q7M487
ID   Q7M487      PRELIMINARY;      PRT;    14 AA.
AC   Q7M487;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hemocyanin chain 3C (Fragment).
OS   Androctonus australis (Sahara scorpion).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthida; Buthoidea; Buthidae; Androctonus.
OX   NCBI_TaxID=6858;
RN   [1]
RP   SEQUENCE.
RA   Jolles J., Jolles P., Lamy J., Lamy J.;
RT   "Structural characterization of seven different subunits in
RT   Androctonus australis haemocyanin.";
RL   FEBS Lett. 106:289-291(1979).
DR   PIR; G61308; G61308.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1670 MW;  7516B699DF604114 CRC64;

  Query Match             27.8%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 2e+04;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RLLMLF 7
              |:| ||
Db          8 RILSLF 13


Search completed: February 22, 2005, 09:37:59
Job time : 53.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-10
Perfect score: 72
Sequence: 1 DRLLMLFAKDVVSRN 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	72	100.0	15	2	AAW04850	Aaw04850 Internal
2	31	43.1	15	7	ADK82698	Adk82698 Beta-amyl
3	30	41.7	15	3	AAB19295	Aab19295 Immunogen
4	29	40.3	10	8	ADK08674	Adk08674 Human pap
5	29	40.3	15	2	AAW97806	Aaw97806 PTEN pept
6	28	38.9	15	6	ABU79062	Abu79062 Aggregati
7	28	38.9	15	7	ABW00196	Abw00196 Peptide #
8	27	37.5	10	4	AAB46229	Aab46229 Human APP
9	27	37.5	10	4	AAG96422	Aag96422 Human com
10	27	37.5	11	4	AAM52586	Aam52586 Peptide #
11	27	37.5	12	2	AAR60372	Aar60372 Beta-amyl
12	27	37.5	12	3	AAB10957	Aab10957 Bovine AD
13	27	37.5	12	8	ADR16408	Adr16408 Human Abe
14	27	37.5	13	5	ABP46621	Abp46621 Human BLy
15	27	37.5	13	6	AAE35465	Aae35465 Abeta pep
16	27	37.5	13	6	AAE35467	Aae35467 Abeta pep
17	27	37.5	13	7	ADG97448	Adg97448 scFV VHCD
18	27	37.5	13	8	ADQ37408	Adq37408 Amyloid-b
19	27	37.5	13	8	ADQ37290	Adq37290 Vaccine a
20	27	37.5	14	2	AAY02792	Aay02792 Fragment
21	27	37.5	14	4	AAE03423	Aae03423 Peptide c
22	27	37.5	14	5	ABP46357	Abp46357 Human BLy
23	27	37.5	14	5	ABP46557	Abp46557 Human BLy
24	27	37.5	14	7	ADA07483	Ada07483 Human sec
25	27	37.5	14	7	ADG97184	Adg97184 scFV VHCD
26	27	37.5	14	7	ADG97384	Adg97384 scFV VHCD
27	27	37.5	14	8	ADN41360	Adn41360 Novel hum
28	27	37.5	14	8	ADN36955	Adn36955 Peptide f
29	27	37.5	15	2	AAW02334	Aaw02334 Beta-amyl
30	27	37.5	15	2	AAW89354	Aaw89354 Beta-amyl
31	27	37.5	15	3	AAY65509	Aay65509 Oestrogen
32	27	37.5	15	5	ABG71014	Abg71014 Long form
33	27	37.5	15	5	ABB05162	Abb05162 Beta amyl
34	27	37.5	15	5	AAU86246	Aau86246 Oestrogen
35	27	37.5	15	5	AAE26271	Aae26271 Human bet
36	27	37.5	15	6	ABU79057	Abu79057 Aggregati
37	27	37.5	15	6	ABU79064	Abu79064 Aggregati
38	27	37.5	15	6	ABU79058	Abu79058 Aggregati
39	27	37.5	15	6	ABU79055	Abu79055 Aggregati
40	27	37.5	15	6	ABU79056	Abu79056 Aggregati
41	27	37.5	15	7	ABW00192	Abw00192 Peptide #
42	27	37.5	15	7	ABW00190	Abw00190 Peptide #
43	27	37.5	15	7	ABW00198	Abw00198 Peptide #
44	27	37.5	15	7	ABW00189	Abw00189 Peptide #
45	27	37.5	15	7	ABW00191	Abw00191 Peptide #

ALIGNMENTS

RESULT 1
AAW04850
ID AAW04850 standard; peptide; 15 AA.
XX
AC AAW04850;
XX
DT 18-FEB-1997 (first entry)
XX
DE Internal fragment of Pseudomonas aeruginosa phosphomannomutase.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Pseudomonas aeruginosa.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 2; Page 44; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is an internal peptide of

```
CC   Pseudomonas aeruginosa phosphomannomutase protein and is implicated as a
CC   foreign epitope involved in the aetiology or in remissions of multiple
CC   sclerosis. It has been shown capable of inducing the proliferation of
CC   autoreactive T-cell clones isolated from multiple sclerosis patients
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 72;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 DRLLMLFAKDVVSRN 15
              ||||||||||||||||
Db          1 DRLLMLFAKDVVSRN 15


RESULT 2
ADK82698
ID   ADK82698 standard; peptide; 15 AA.
XX
AC   ADK82698;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Beta-amyloid peptide #4 recognised by antibody to treat senile dementia.
XX
KW   fusion antibody; senile dementia; beta-amyloid peptide; fibre;
KW   immunocell.
XX
OS   Homo sapiens.
XX
PN   CN1396183-A.
XX
PD   12-FEB-2003.
XX
PF   13-JUL-2001; 2001CN-00120278.
XX
PR   13-JUL-2001; 2001CN-00120278.
XX
PA   (ZHAN/) ZHANG X.
XX
PI   Zhang X,  Zhang J;
XX
DR   WPI; 2003-442233/42.
XX
PT   Human fusion antibody for reducing cerebral amyloid fibers associated
PT   with senile dementia.
XX
PS   Claim 1; Page 2; 26pp; Chinese.
XX
CC   The invention relates to a human fusion antibody for preventing and
CC   treating senile dementia. The antibody recognises and binds the beta-
CC   amyloid peptide and the fibres generated by it. The human antibody Fc
CC   segment recognized by human immunocells are sequentially contained by its
CC   terminals from N to C. The fusion gene coding for the antibody is also
CC   disclosed. This sequence represents a beta-amyloid peptide recognised by
CC   the antibody.
XX
SQ   Sequence 15 AA;

  Query Match             43.1%;  Score 31;  DB 7;  Length 15;
  Best Local Similarity   70.0%;  Pred. No. 81;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 LMLFAKDVVS 13
              |: |||||  |
Db          2 LVFFAKDVGS 11


RESULT 3
AAB19295
ID   AAB19295 standard; peptide; 15 AA.
XX
AC   AAB19295;
XX
DT   19-FEB-2001  (first entry)
XX
DE   Immunogenic peptide useful for generating antigen specific T cells.
XX
KW   Immunogenic peptide; antigen specific T cell; cell proliferation; cancer;
KW   inflammation; autoimmune disease; dermatological disorder;
KW   neurodegenerative disorder; atherosclerosis; rheumatoid arthritis;
KW   osteoporosis; chronic ulcer; psoriasis; cardiovascular disease;
KW   infectious disease.
XX
OS   Physeter catodon.
XX
PN   WO200057920-A2.
XX
PD   05-OCT-2000.
XX
PF   30-MAR-2000; 2000WO-GB001225.
XX
PR   30-MAR-1999;   99GB-00007366.
XX
PA   (MEDI-) MEDICAL RES COUNCIL.
XX
PI   Fisher A;
XX
DR   WPI; 2000-647208/62.
XX
PT   Regulating expression of transgenes encoding enzymes or hormones in
PT   mammals, by transforming expandable population of cells with desired
PT   transgene and regulating cell proliferation by administration of an
PT   agent.
XX
PS   Disclosure; Page 22; 37pp; English.
XX
CC   The present sequence represents an immunogenic peptide, derived from
CC   myoglobin, which is useful for generating antigen specific T cells. It
CC   may be used in the course of the invention. The specification describes
CC   the use of an agent which regulates cell proliferation for modulating the
CC   levels of production of a gene product of interest in a host organism.
CC   The method comprises transforming an expandable population of cells with
CC   a transgene encoding the gene product, expressing the gene in the host
CC   and regulating proliferation of the population of cells by administration
CC   of the agent. The method is useful for regulating cell proliferation. The
CC   method is also useful for delivering one or more transgenes useful in the
CC   treatment of cancer, inflammation, autoimmune disease, dermatological
CC   disorder, neurodegenerative disorder, atherosclerosis, rheumatoid
CC   arthritis, osteoporosis, chronic ulcers, psoriasis, cardiovascular
CC   diseases and infectious diseases
XX
SQ   Sequence 15 AA;

  Query Match             41.7%;  Score 30;  DB 3;  Length 15;
  Best Local Similarity   35.7%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 DRLLMLFAKDVVSR 14
              :: | || ||: ::
Db          1 NKALELFRKDIAAK 14


RESULT 4
ADK08674
ID   ADK08674 standard; peptide; 10 AA.
XX
AC   ADK08674;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human papillomavirus peptide #729.
XX
KW   pathogenic virus; alternative reading frame; antigenic determinant;
KW   virucide; vaccine; therapeutic agent; infection; HPV.
```

XX
OS Human papillomavirus.
XX
PN WO2004011650-A2.
XX
PD 05-FEB-2004.
XX
PF 24-JUL-2003; 2003WO-EP008112.
XX
PR 24-JUL-2002; 2002AT-00001124.
PR 11-JUL-2003; 2003EP-00450171.
XX
PA (INTE-) INTERCELL AG.
XX
PI Mattner F, Schmidt W, Habel A;
XX
DR WPI; 2004-169243/16.
XX
PT New polypeptide encoded by an alternative reading frame of a pathogenic
PT virus comprising an antigenic determinant, useful for treating or
PT preventing an infection with the pathogenic virus.
XX
PS Claim 18; Page 181; 220pp; English.
XX
CC This invention relates to a novel polypeptide encoded by an alternative
CC reading frame of a pathogenic virus, where the polypeptide starts with a
CC methionine amino acid residue, which comprises an antigenic determinant
CC and more than 7 amino acid residues. The invention may be useful for the
CC production of compounds with a virucide activity or the development of a
CC vaccine. The polypeptide or its fragments may be useful as a therapeutic
CC agent. It is also useful for the manufacture of a medicament for treating
CC or preventing an infection with the pathogenic virus. The present
CC sequence is that of a human papillomavirus (HPV) epitope peptide of the
CC invention.
XX
SQ Sequence 10 AA;

Query Match 40.3%; Score 29; DB 8; Length 10;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 MLPAKDVV 12
:|| ||||
Db 1 ILPLKDVV 8

RESULT 5
AAW97806
ID AAW97806 standard; peptide; 15 AA.
XX
AC AAW97806;
XX
DT 21-MAY-1999 (first entry)
XX
DE PTEN peptide (aa6-19) used for antibody generation.
XX
KW PTEN; MMAC1; protein tyrosine phosphatase; human; prostate cancer;
KW brain cancer; prostate cancer; tumour suppressor; Cowden's disease;
KW neurodegenerative disease; Parkinson's disease; diagnosis; therapy;
KW antibody.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9902704-A2.
XX
PD 21-JAN-1999.
XX
PF 08-JUL-1998; 98WO-US014205.
XX
PR 08-JUL-1997; 97US-0051908P.
PR 29-JUN-1998; 98US-0090984P.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
XX
PI Tonks NK, Myers MP;
XX
DR WPI; 1999-120905/10.
XX
PT New use of PTEN phosphatase - for developing products for the diagnosis
PT and treatment of hyperproliferative disorders, e.g. cancers or
PT neurodegenerative disorders such as Parkinson's disease.
XX
PS Example 2; Page 35; 60pp; English.
XX
CC This peptide corresponds to amino acids Lys6-Asp19 of PTEN (see
CC AAW97802), a novel human dual specificity phosphatase that has tumour
CC suppressor activity. It also includes a C-terminal Cys residue, not found
CC in the wild-type PTEN, which was added to aid in the conjugation of the
CC peptide to keyhole limpet haemocyanin. PTEN peptides (see AAW97805-08)
CC were used to immunise rabbits. A rabbit injected with PTEN Glu388-Val403
CC produced antibodies that recognised endogenous PTEN on immunoblots and in
CC immunoprecipitations; this antibody was termed PTEN486. All other rabbits
CC failed to produce antibodies that could recognise PTEN. PTEN polypeptides
CC and polynucleotides (see AAX07339) can be used to develop products for
CC the diagnosis and treatment of hyperproliferative disorders such as
CC cancers, or neurodegenerative disorders such as Parkinson's disease
XX
SQ Sequence 15 AA;

Query Match 40.3%; Score 29; DB 2; Length 15;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 9 KDVVSRN 15
|::||||
Db 1 KEIVSRN 7

RESULT 6
ABU79062
ID ABU79062 standard; peptide; 15 AA.
XX
AC ABU79062;
XX
DT 17-JUN-2003 (first entry)
XX
DE Aggregation blocking peptide #14.
XX
KW Amyloid formation; amyloid-like deposit; Alzheimer's disease;
KW pathological beta-sheet-rich conformation; Down's syndrome;
KW amyloidosis disorder; human prion disease; kuru; CJD;
KW Creutzfeldt-Jakob disease; Gerstmann-Straussler-Scheinker syndrome; GSS;
KW prion associated human neurodegenerative disease; animal prion disease;
KW scrapie; spongiform encephalopathy; transmissible mink encephalopathy;
KW chronic wasting disease.
XX
OS Unidentified.
XX
PN US6462171-B1.
XX
PD 08-OCT-2002.
XX
PF 12-DEC-1996; 96US-00766596.
XX
PR 07-JUN-1995; 95US-00478326.
PR 10-APR-1996; 96US-00630645.
XX
PA (UYNY) UNIV NEW YORK STATE.
XX
PI Soto-Jara C, Baumann MH, Frangione B;
XX
DR WPI; 2003-379012/36.
XX
PT Novel inhibitory peptides which inhibit and structurally block abnormal
PT folding of protein into amyloid or amyloid-like deposit and into

PT pathological beta-sheet rich conformation, useful for treating
PT Alzheimer's disease.
XX
PS Disclosure; Col 51-52; 51pp; English.
XX
CC The invention describes an isolated inhibitory peptide (I) which
CC interacts with a hydrophobic beta-sheet forming cluster of amino acid
CC residues on a protein or peptide for amyloid or amyloid-like deposit
CC formation, and inhibits or structurally blocks the abnormal folding of
CC proteins and peptides into amyloid or amyloid-like deposits and into
CC pathological beta-sheet-rich conformation. (I) is useful for disorders or
CC diseases associated with abnormal protein folding into amyloid or amyloid
CC -like deposits or into pathological beta-sheet-rich precursors of such
CC deposits, such as Alzheimer's disease, Down's syndrome, other amyloidosis
CC disorders, human prion diseases, such as kuru, Creutzfeldt-Jakob disease
CC (CJD), Gerstmann-Straussler-Scheinker syndrome (GSS), prion associated
CC human neurodegenerative diseases as well as animal prion diseases such as
CC scrapie, spongiform encephalopathy, transmissible mink encephalopathy and
CC chronic wasting disease of mule deer and elk. (I) is also useful for
CC detecting and diagnosing the presence or absence of amyloid or amyloid-
CC like deposits in vivo and its precursors. This is the amino acid sequence
CC of peptide associated with the inhibition of amyloid or amyloid like
CC deposits
XX
SQ Sequence 15 AA;

Query Match 38.9%; Score 28; DB 6; Length 15;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LMLFAKDVVS 13
|: ||:|| |
Db 6 LVFFAEDVPS 15

RESULT 7
ABW00196
ID ABW00196 standard; peptide; 15 AA.
XX
AC ABW00196;
XX
DT 15-JAN-2004 (first entry)
XX
DE Peptide #14 used in the invention.
XX
KW Amyloid-like fibril deposit; prion related encephalopathy; gene therapy;
KW Alzheimer's disease.
XX
OS Unidentified.
XX
PN US2003087407-A1.
XX
PD 08-MAY-2003.
XX
PF 06-SEP-2002; 2002US-00235483.
XX
PR 07-JUN-1995; 95US-00478326.
PR 10-APR-1996; 96US-00630645.
PR 12-DEC-1996; 96US-00766596.
XX
PA (UYNY) UNIV NEW YORK STATE.
XX
PI Soto-Jara C, Baumann MH, Frangione B;
XX
DR WPI; 2003-616149/58.
XX
PT New inhibitory peptide, useful for preparing a composition for
PT diagnosing, preventing or treating disorders associated with amyloid-like
PT fibril deposits, e.g. Alzheimer's disease, or prion related
PT encephalopathies.
XX
PS Claim 1; Page 27; 52pp; English.
XX
CC The invention relates to inhibitory peptide comprising a portion of at
CC least three amino acid residues and a sequence predicted not to adopt a
CC beta-sheet structure that associates with a hydrophobic beta-sheet
CC cluster on a protein or peptide involved in the abnormal folding into a
CC beta-sheet structure, to structurally block the abnormal folding of the
CC protein or peptide. The inhibitory peptide is useful for preparing a
CC composition for preventing, treating or detecting disorders or diseases
CC associated with amyloid-like fibril deposits e.g. Alzheimer's disease and
CC prion related encephalopathies. The invention is also useful in gene
CC therapy. The present sequence is a peptide used in the invention
XX
SQ Sequence 15 AA;

Query Match 38.9%; Score 28; DB 7; Length 15;
Best Local Similarity 60.0%; Pred. No. 3e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 LMLFAKDVVS 13
|: ||:|| |
Db 6 LVFFAEDVPS 15

RESULT 8
AAB46229
ID AAB46229 standard; peptide; 10 AA.
XX
AC AAB46229;
XX
DT 04-APR-2001 (first entry)
XX
DE Human APP derived immunogenic peptide #25.
XX
KW Amyloid deposit; APP; Abeta; brain; human; clearing response; nootropic;
KW Fc receptor mediated phagocytosis; immunogenic response; neuroprotective;
KW amyloid precursor protein; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO200072880-A2.
XX
PD 07-DEC-2000.
XX
PF 26-MAY-2000; 2000WO-US014810.
XX
PR 28-MAY-1999; 99US-00322289.
XX
PA (NEUR-) NEURALAB LTD.
XX
PI Schenk DB, Bard F, Vasquez NJ, Yednock T;
XX
DR WPI; 2001-032104/04.
XX
PT Preventing or treating a disease associated with amyloid deposits,
PT especially Alzheimer's disease, comprises administering amyloid specific
PT antibody.
XX
PS Disclosure; Fig 19; 143pp; English.
XX
CC This invention describes a novel method of preventing or treating a
CC disease associated with amyloid deposits of amyloid precursor protein
CC (APP) Abeta fragments in the brain of a patient, which comprises
CC administering to the patient: (a) an antibody that binds to Abeta, the
CC antibody binds to an amyloid deposit and induces a clearing response (Fc
CC receptor mediated phagocytosis) against it (b) a polypeptide containing
CC an N-terminal segment of at least residues 1-5 of Abeta; or (c) an agent
CC that induces an immunogenic response against residues 1-3 to 7-11 of
CC Abeta. The products of the invention have nootropic and neuroprotective
CC activity. The method is also useful for monitoring a course of treatment
CC being administered to a patient e.g. active and passive immunization. The
CC methods are useful for prophylactic and therapeutic treatment of
CC Alzheimer's disease
XX
SQ Sequence 10 AA;

```
  Query Match             37.5%;  Score 27;  DB 4;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 2.9e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 LMLFAKDVVS 13
              |: ||:|| |
Db          1 LVFFAEDVGS 10


RESULT 9
AAG96422
ID   AAG96422 standard; peptide; 10 AA.
XX
AC   AAG96422;
XX
DT   18-SEP-2001  (first entry)
XX
DE   Human complementary peptide, SEQ ID NO: 2616.
XX
KW   Human; complementary peptide; ligand; drug discovery; drug design.
XX
OS   Homo sapiens.
XX
PN   WO200142277-A2.
XX
PD   14-JUN-2001.
XX
PF   13-DEC-2000; 2000WO-GB004776.
XX
PR   13-DEC-1999;   99GB-00029464.
XX
PA   (PROT-) PROTEOM LTD.
XX
PI   Roberts GW,  Heal JR;
XX
DR   WPI; 2001-408419/43.
XX
PT   A set of peptide ligands consisting of specific complementary peptides to
PT   proteins encoded by genes of the human genome, useful in an assay for
PT   screening and identifying of one or more novel peptides which are drug
PT   candidates or pro-drugs.
XX
PS   Example 4; Page 416; 646pp; English.
XX
CC   The invention relates to a set of complementary peptide ligands generated
CC   from the human genome. The complementary peptides interact with their
CC   relevant target proteins encoded in the human genome. They can be used as
CC   reagents in drug discovery and as lead ligands to facilitate drug design
CC   and development. The present sequence is a complementary peptide provided
CC   in the specification
XX
SQ   Sequence 10 AA;

  Query Match             37.5%;  Score 27;  DB 4;  Length 10;
  Best Local Similarity   44.4%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 FAKDVVSRN 15
              ::|| : ||
Db          2 YSKDTLGRN 10


RESULT 10
AAM52586
ID   AAM52586 standard; peptide; 11 AA.
XX
AC   AAM52586;
XX
DT   07-FEB-2002  (first entry)
XX
DE   Peptide #16 for illustrating method of anticipating protein interaction.
XX
KW   Protein interaction; biochemistry; molecular biology; drug development;
KW   agrochemical; bioengineering.
XX
OS   Unidentified.
XX
PN   WO200167299-A1.
XX
PD   13-SEP-2001.
XX
PF   09-MAR-2001; 2001WO-JP001846.
XX
PR   10-MAR-2000; 2000JP-00072485.
XX
PA   (DAUC ) DAIICHI PHARM CO LTD.
PA   (FUIT ) FUJITSU LTD.
XX
PI   Doi H,  Suzuki A;
XX
DR   WPI; 2001-570799/64.
XX
PT   Method for assaying a specific protein for assaying anticipated
PT   information.
XX
PS   Example 14; Page 34; 64pp; Japanese.
XX
CC   The present invention relates to a method for anticipating interaction
CC   between proteins. The method comprises (1) digesting protein A into
CC   oligopeptides; (2) searching a protein sequence database for polypeptides
CC   (polypeptide C) containing these oligopeptide sequences or D their
CC   homologues; (3) performing a local alignment of A and detected C or D;
CC   and (4) using a value calculated from the amino acid or oligonucleotide
CC   frequencies, anticipating that C or D is polypeptide B that interacts
CC   with A. The method is useful for assaying anticipated information about
CC   proteins in biochemical, molecular biology, drug development,
CC   agrochemical and bioengineering areas. The present sequence was used to
CC   illustrate the method
XX
SQ   Sequence 11 AA;

  Query Match             37.5%;  Score 27;  DB 4;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 3.3e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 LMLFAKDVVS 13
              |: ||:|| |
Db          2 LVFFAEDVGS 11


RESULT 11
AAR60372
ID   AAR60372 standard; peptide; 12 AA.
XX
AC   AAR60372;
XX
DT   25-MAR-2003  (revised)
DT   15-MAR-1995  (first entry)
XX
DE   Beta-amyloid (17-28).
XX
KW   Amyloid precursor protein; APP; Alzheimer's disease; beta-amyloid;
KW   anti-beta-amyloid antibody; diagnosis.
XX
OS   Homo sapiens.
XX
PN   WO9417197-A1.
XX
PD   04-AUG-1994.
XX
PF   24-JAN-1994;   94WO-JP000089.
XX
PR   25-JAN-1993;   93JP-00010132.
PR   05-FEB-1993;   93JP-00019035.
PR   16-NOV-1993;   93JP-00286985.
```

```
PR   28-DEC-1993;   93JP-00334773.
XX
PA   (TAKE ) TAKEDA CHEM IND LTD.
XX
PI   Suzuki N,  Odaka A,  Kitada C;
XX
DR   WPI; 1994-264110/32.
XX
PT   Antibodies recognising specific parts of beta-amyloid - can be used for
PT   diagnosis of diseases implicating beta-amyloid, such as Alzheimer's
PT   disease.
XX
PS   Disclosure; Page 85; 116pp; Japanese.
XX
CC   Antibodies which recognise specific subfragments of the beta-amyloid
CC   protein are claimed. Specifically, the antibodies (which are pref.
CC   monoclonal) recognise residues 1-16 and/or 1-28 from the N-terminal
CC   portion of beta-amyloid or they recognise residues 25-35 or 35-43 from
CC   the C-terminal portion. The antibodies are useful for assaying beta-
CC   amyloid and its derivatives for diagnosis of Alzheimer's disease.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 12 AA;

  Query Match             37.5%;  Score 27;  DB 2;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.6e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 LMLFAKDVVS 13
              |: ||:|| |
Db          1 LVFFAEDVGS 10


RESULT 12
AAB10957
ID   AAB10957 standard; protein; 12 AA.
XX
AC   AAB10957;
XX
DT   07-FEB-2001  (first entry)
XX
DE   Bovine ADAM 10 peptide substrate #2.
XX
KW   APP; amyloid precursor protein; human; alpha-secretase; ADAM 10;
KW   disintegrin-metalloprotease; protease; nootropic; neuroprotective;
KW   gene therapy; Alzheimer's disease.
XX
OS   Bos taurus.
XX
PN   DE19910108-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-1999;   99DE-01010108.
XX
PR   08-MAR-1999;   99DE-01010108.
XX
PA   (FAHR/) FAHRENHOLZ F.
XX
PI   Fahrenholz F,  Postina R;
XX
DR   WPI; 2000-588391/56.
XX
PT   Recombinant cells, for identifying alpha-secretase active agents and
PT   identifying risk factors associated with Alzheimer's disease, comprise
PT   amyloid precursor protein and alpha-secretase.
XX
PS   Disclosure; Page 6; 24pp; German.
XX
CC   This invention describes a novel recombinant cell comprising recombinant
CC   nucleic acids encoding a region of human amyloid precursor protein
CC   containing an alpha-secretase cleavage site and a protease or a
CC   heterologous RNA coding for a substrate protein and a protease. The
CC   invention also describes a recombinant cell, characterized in that it
CC   contains recombinant nucleic acids comprising either: (a) a gene for a
CC   substrate protein (SP), which comprises a sequence region of 18 amino
CC   acids of the human amyloid precursor protein (APP) or a homologous
CC   protein, where the sequence region contains the alpha-secretase cleavage
CC   site at a reference of 6 residues at the N-terminal and 12 residues at
CC   the C-terminal; and (b) a gene for a protease protein (PP), that either
CC   comprises a proteolytically active necessary sequence region or a
CC   sequence region of the disintegrin metalloprotease ADAM 10 from a cow
CC   (Bos taurus), from a human or other mammal or a mutant of this, which
CC   shows the same enzymatic properties, where the genes are under the
CC   control of heterologous promoters; or a heterologous RNA coding for a SP
CC   and a PP. The products of the invention have nootropic and
CC   neuroprotective activity and can be used for gene therapy. The protease
CC   proteins of the invention are useful for proteolytic cleavage of
CC   substrate proteins, especially human amyloid precursor protein. Dominant
CC   negative forms of bovine, human or other mammalian disintegrin-
CC   metalloprotease ADAM 10 proteins and their coding sequences are useful
CC   for suppressing the alpha-secretase activity of a cell. Nucleic acid
CC   sequences encoding the proteases are useful for constructing vectors for
CC   gene therapy. The proteins and recombinant cells are useful for
CC   identifying secretases and pharmaceutical agents and to identify risk
CC   factors associated with Alzheimer's disease
XX
SQ   Sequence 12 AA;

  Query Match             37.5%;  Score 27;  DB 3;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.6e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 LMLFAKDVVS 13
              |: ||:|| |
Db          1 LVFFAEDVGS 10


RESULT 13
ADR16408
ID   ADR16408 standard; peptide; 12 AA.
XX
AC   ADR16408;
XX
DT   21-OCT-2004  (first entry)
XX
DE   Human Abeta (amyloid-beta) tryptic peptide.
XX
KW   Alzheimer's Disease; AD; occular tissue; lens; amyloidogenic disorder;
KW   Creutzfeld-Jakob disease; spongiform encephalopathy; Prion disease;
KW   scrapie; bovine spongiform encephalopathy; veterinary prionopathy;
KW   Parkinson's disease; Huntington's disease; trinucleotide repeat disease;
KW   amyotrophic lateral sclerosis; Down's Syndrome; Pick's Disease;
KW   Frontotemporal Dementia; Lewy Body Disease; Hallervorden-Spatz Disease;
KW   synucleinopathy; multiple system atrophy;
KW   neuronal intranuclear inclusion disease; tauopathy;
KW   progressive supranuclear palsy; corticobasal degeneration; human;
KW   amyloid-beta; Abeta.
XX
OS   Homo sapiens.
XX
PN   US2004152068-A1.
XX
PD   05-AUG-2004.
XX
PF   18-NOV-2003; 2003US-00715776.
XX
PR   21-AUG-2000; 2000US-0226590P.
PR   27-APR-2001; 2001US-0287124P.
PR   21-AUG-2001; 2001US-00935126.
PR   25-APR-2002; 2002US-00132779.
PR   18-NOV-2002; 2002US-0427153P.
PR   05-MAR-2003; 2003US-0452336P.
XX
PA   (GOLD/) GOLDSTEIN L E.
PA   (CHYL/) CHYLACK L T.
```

```
XX
PI   Goldstein LE,  Chylack LT;
XX
DR   WPI; 2004-580178/56.
XX
PT   Method for diagnosing Alzheimer's disease in a mammal involves contacting
PT   an ocular tissue with a detectably-labeled compound that binds to an
PT   amyloid protein or use of magnetic resonance imaging.
XX
PS   Example 3; SEQ ID NO 1; 23pp; English.
XX
CC   The invention relates to a method for diagnosing or providing a prognosis
CC   regarding the state of Alzheimer's Disease (AD) in a mammal. The method
CC   involves contacting an ocular tissue with a detectably-labelled compound
CC   that binds to an amyloid protein where an increase in binding of the
CC   compound to the ocular tissue compared to the normal control level of
CC   binding indicates that the mammal is suffering from or is at risk of
CC   developing AD. The method is also used for the diagnosis of amyloidogenic
CC   disorder e.g. Familial AD, Sporadic AD, Creutzfeld-Jakob disease, variant
CC   Creutzfeld-Jakob disease, spongiform encephalopathies, Prion diseases
CC   (including scrapie, bovine spongiform encephalopathy and other veterinary
CC   prionopathies), Parkinson's disease, Huntington's disease (and
CC   trinucleotide repeat diseases), amyotrophic lateral sclerosis, Down's
CC   Syndrome (Trisomy 21), Pick's Disease (Frontotemporal Dementia), Lewy
CC   Body Disease, neurodegeneration with brain iron accumulation
CC   (Hallervorden-Spatz Disease), synucleinopathies (including Parkinson's
CC   disease, multiple system atrophy, dementia with Lewy Bodies, and others),
CC   neuronal intranuclear inclusion disease, tauopathies (including
CC   progressive supranuclear palsy, Pick's disease, corticobasal
CC   degeneration, hereditary frontotemporal dementia (with or without
CC   Parkinsonism) and Guam amyotrophic lateral sclerosis/parkinsonism
CC   dementia complex). The present sequence is human amyloid-beta (Abeta)
CC   peptide. This sequence is used to illustrate the method of the invention.
XX
SQ   Sequence 12 AA;

  Query Match             37.5%;  Score 27;  DB 8;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 3.6e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 LMLFAKDVVS 13
              |: ||:|| |
Db          1 LVFFAEDVGS 10


RESULT 14
ABP46621
ID   ABP46621 standard; peptide; 13 AA.
XX
AC   ABP46621;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human BLyS binding scFv VH CDR3 SEQ ID 2632.
XX
KW   BLys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW   tumour necrosis factor; B cell proliferation; B cell differentiation;
KW   immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW   antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW   systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW   common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS   Homo sapiens.
XX
PN   WO200202641-A1.
XX
PD   10-JAN-2002.
XX
PF   15-JUN-2001; 2001WO-US019110.
XX
PR   16-JUN-2000; 2000US-0212210P.
PR   17-OCT-2000; 2000US-0240816P.
PR   16-MAR-2001; 2001US-0276248P.
PR   21-MAR-2001; 2001US-0277379P.
PR   25-MAY-2001; 2001US-0293499P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI   Ruben SM,  Barash SC,  Choi GH,  Vaughan T,  Hilbert D;
XX
DR   WPI; 2002-114799/15.
XX
PT   Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
PT   diagnosis and treatment of cancers and immune disorders.
XX
PS   Claim 2; Page 3023; 3148pp; English.
XX
CC   This invention describes novel antibodies that immunospecifically bind to
CC   B Lymphocyte Stimulator (BLyS) polypeptides. BLyS is a member of the
CC   tumour necrosis factor (TNF) super family and induces B cell
CC   proliferation and differentiation. The antibodies of the invention have
CC   cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC   antirheumatic and antiAIDS activity and can be used in vaccines to
CC   inhibit the expression and activity of BLyS. The antibodies bind to BLyS
CC   and so may be used to detect and quantitate the presence of BLyS in
CC   biological samples and may be used in this way to diagnose disease
CC   associated with aberrant expression of BLyS. They may also be
CC   administered to treat diseases associated with aberrant BLyS expression
CC   and activity such as cancer, immune, and autoimmune disorders and
CC   diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC   immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC   acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC   the antibodies and fragments of the antibodies described in the method of
CC   the invention
XX
SQ   Sequence 13 AA;

  Query Match             37.5%;  Score 27;  DB 5;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 3.9e+02;
  Matches    6;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RLLMLFAKDVVS 13
              | |:||  | :|
Db          2 RDLLLFPNDALS 13


RESULT 15
AAE35465
ID   AAE35465 standard; peptide; 13 AA.
XX
AC   AAE35465;
XX
DT   17-JUN-2003  (first entry)
XX
DE   Abeta peptide #36.
XX
KW   All-D-amyloid-beta peptide; Alzheimer's disease; rheumatoid arthritis;
KW   cerebral amyloid angiopathy; amyloid disease; ankylosing spondylitis;
KW   psoriasis; Reiter's syndrome; Adult Still's disease; Bechet's syndrome;
KW   Crohn's disease; infection; leprosy; tuberculosis; carcinoma; nootropic;
KW   chronic pyelonephritis; osteomyelitis; Whipple's disease; vasotropic;
KW   Hodgkin's lymphoma; neuroprotective; bronchiectasis; ophthalmological;
KW   ulcer; antiinflammatory; cytostatic; uropathic; therapy.
XX
OS   Unidentified.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1. .6
FT                   /note= "D-form residues"
XX
PN   WO200296937-A2.
XX
PD   05-DEC-2002.
XX
PF   29-MAY-2002; 2002WO-CA000763.
```

XX
PR 29-MAY-2001; 2001US-00867847.
XX
PA (NEUR-) NEUROCHEM INC.
XX
PI Gervais F, Hebert L, Chalifour RJ, Kong X;
XX
DR WPI; 2003-201269/19.
XX
PT Prevention and/or treatment of an amyloid-related disease e.g.
PT Alzheimer's disease, comprises use of all-D-amyloid-beta peptides.
XX
PS Claim 1; Page 61; 44pp; English.
XX
CC The invention relates to a method for prevention and/or treatment of an
CC amyloid-related disease which comprises administration of an all-D -
CC amyloid-beta peptide. The method is used for preventing and/or treating
CC Alzheimer's and other amyloid related disease e.g. cerebral amyloid
CC angiopathy; for altering serum levels of amyloid-beta in a mammal and
CC favours the clearance of soluble amyloid-beta or fibril amyloid-beta from
CC the mammal; and reducing or inhibiting the formation of plaques. It is
CC also used for treating AA (reactive) amyloid diseases including
CC inflammatory diseases e.g. rheumatoid arthritis, juvenile chronic
CC arthritis, ankylosing spondylitis, psoriasis, psoriatic arthropathy,
CC Reiter's syndrome, Adult Still's disease, Bechet's syndrome and Crohn's
CC disease. AA deposits are also produced as a result of chronic microbial
CC infections (preferably leprosy, tuberculosis, bronchiectasis, decubitus
CC ulcers, chronic pyelonephritis, osteomyelitis and Whipple's disease).
CC Certain malignant neoplasms can also result in AA fibril amyloid deposits
CC including Hodgkin's lymphoma, renal carcinoma, carcinomas of gut, lung
CC and urogenital tract, basal cell carcinoma and hairy cell leukaemia. The
CC present sequence is an Abeta peptide used to illustrate the method of the
CC invention
XX
SQ Sequence 13 AA;

Query Match 37.5%; Score 27; DB 6; Length 13;
Best Local Similarity 60.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 LMLFAKDVVS 13
              |: ||:|| |
Db          2 LVFFAEDVGS 11
```

Search completed: February 22, 2005, 09:24:48
Job time : 67.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-11
Perfect score: 79
Sequence: 1 IGGRVHFFKDISPIA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	26	32.9	15	2	PH0775	T-cell receptor al
2	26	32.9	15	2	PH0779	T-cell receptor al
3	26	32.9	15	2	PH1455	T-cell receptor al
4	24	30.4	12	2	S26557	T-cell receptor be
5	23	29.1	14	2	S29632	xylan 1,4-beta-xyl
6	22	27.8	10	2	A58365	neuropeptide FFRFa
7	22	27.8	11	2	S60354	retinal oxidase -
8	21	26.6	10	2	A60476	S-layer protein -
9	21	26.6	11	2	PC2330	cycloinulooligosac
10	21	26.6	12	2	S01749	collagen alpha 1(I
11	21	26.6	14	2	PN0151	omega-gliadine 2'
12	20	25.3	9	2	PT0670	T-cell receptor be
13	20	25.3	13	2	S15755	actin 7 - soybean
14	20	25.3	14	2	PH0804	T-cell receptor al
15	20	25.3	14	2	PC4376	telomeric and tetr
16	20	25.3	15	2	S24159	leukocyte elastase
17	20	25.3	15	2	PA0054	protein QF200017 -
18	20	25.3	15	2	PA0061	protein QF200039 -
19	19	24.1	7	2	PN0150	omega-gliadine 1'
20	19	24.1	11	2	A34243	H-hyosophorin - Ja
21	19	24.1	12	2	A09985	gamma-crystallin -
22	19	24.1	12	2	PT0228	Ig heavy chain CDR
23	19	24.1	13	2	D61491	seed protein ws-4
24	19	24.1	14	2	F33160	H+-transporting tw
25	19	24.1	14	2	PN0147	omega-gliadine 1 a
26	19	24.1	15	2	PT0205	insulin-like growt
27	19	24.1	15	2	PD0444	coupling factor 6
28	18	22.8	11	2	PC2173	triacylglycerol li
29	18	22.8	11	2	S57575	T cell receptor V-
30	18	22.8	12	2	PQ0730	unidentified 5.4/3
31	18	22.8	13	2	S47380	T-cell antigen rec
32	18	22.8	13	2	H56046	urinary tract ston
33	18	22.8	14	2	PH0762	T-cell receptor be
34	18	22.8	15	2	PS0185	27K protein A 3.4/
35	18	22.8	15	2	S29485	GTP-binding protei
36	18	22.8	15	2	S47387	T-cell antigen rec
37	18	22.8	15	2	A61247	urogenital tumor m
38	18	22.8	15	2	G49732	NADH2 dehydrogenas
39	17	21.5	4	2	JQ1273	neuropeptide Antho
40	17	21.5	7	2	I48086	DNA topoisomerase
41	17	21.5	7	2	PT0667	T-cell receptor be
42	17	21.5	9	2	D24180	fibrinogen beta ch
43	17	21.5	9	2	F28854	fibrinopeptide B -
44	17	21.5	9	2	PT0634	T-cell receptor be
45	17	21.5	9	2	A42266	peptidylglycine mo

ALIGNMENTS

RESULT 1
PH0775
T-cell receptor alpha chain (B28) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0775
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0775
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60871
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 32.9%; Score 26; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GRVHF 7
||:||
Db 10 GRLHF 14

RESULT 2
PH0779
T-cell receptor alpha chain (B83) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0779
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0779
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60877
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 32.9%; Score 26; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GRVHF 7
||:||
Db 10 GRLHF 14

RESULT 3
PH1455
T-cell receptor alpha chain (clone A24/PEF4) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Mar-2004
C;Accession: PH1455
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Kc
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompatib
A;Reference number: PH1430; MUID:93171821; PMID:8436911
A;Accession: PH1455
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Keywords: receptor; T-cell

Query Match 32.9%; Score 26; DB 2; Length 15;
Best Local Similarity 80.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GRVHF 7
||:||
Db 10 GRLHF 14

RESULT 4
S26557
T-cell receptor beta chain (clone Cw3/HLA2A3) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C;Accession: S26557
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wid
J. Exp. Med. 176, 439-447, 1992
A;Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor
A;Reference number: S26512; MUID:92364546; PMID:1380061
A;Accession: S26557
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Cross-references: EMBL:X68007
A;Experimental source: cytolytic T-lymphocyte, clone Cw3/HLA2A3
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 30.4%; Score 24; DB 2; Length 12;
Best Local Similarity 66.7%; Pred. No. 5.3e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GRVHFF 8
||| :|
Db 6 GRVEYF 11

RESULT 5
S29632
xylan 1,4-beta-xylosidase (EC 3.2.1.37) - Thermotoga sp. (strain FjSS3-B.1) (fragment)
N;Alternate names: beta-xylosidase
C;Species: Thermotoga sp.
A;Variety: FjSS3-B.1
C;Date: 19-Mar-1997 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: S29632
R;Ruttersmith, L.D.; Daniel, R.M.
Biochim. Biophys. Acta 1156, 167-172, 1993
A;Title: Thermostable beta-glucosidase and beta-xylosidase from Thermotoga sp. strain Fj
A;Reference number: S29631; MUID:93152594; PMID:8427876
A;Accession: S29632
A;Molecule type: protein
A;Residues: 1-14 <RUT>
A;Cross-references: UNIPROT:Q7M0Q6
A;Experimental source: strain FjSS3-B.1
C;Comment: Although the beta-xylosidase enzyme activity was apparently confirmed for thi
C;Function:
A;Description: hydrolyzes short chain oligosacchrides and xylobiose to produce D-xylose
A;Note: plays an important role in the relief of end-product inhibition of endoxylanase
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 29.1%; Score 23; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 9.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 VHFFKD 10
|:|| |
Db 5 VYFFAD 10

RESULT 6
A58365
neuropeptide FFRFamide - blue mussel
N;Alternate names: FMRFamide-related decapeptide; Mytilus FFRFamide
C;Species: Mytilus edulis (blue mussel)
C;Date: 20-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 09-Jul-2004
C;Accession: A58365
R;Fujisawa, Y.; Ikeda, T.; Nomoto, K.; Yasuda-Kamatani, Y.; Minakata, H.; Kenny, P.T.M.;
Comp. Biochem. Physiol. C 102, 91-95, 1992
A;Title: The FMRFamide-related decapeptide of Mytilus contains a D-amino acid residue.
A;Reference number: A58365; MUID:93047882; PMID:1358533
A;Accession: A58365
A;Molecule type: protein
A;Residues: 1-10 <FUJ>
A;Cross-references: UNIPROT:P42560
A;Experimental source: anterior byssus retractor muscle
C;Keywords: amidated carboxyl end; D-amino acid; neuropeptide
F;2/Modified site: D-leucine (Leu) #status experimental
F;10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 27.8%; Score 22; DB 2; Length 10;
Best Local Similarity 75.0%; Pred. No. 9.9e+02;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 HFFK 9
|||:
Db 6 HFFR 9

RESULT 7
S60354
retinal oxidase - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C;Accession: S60354
R;Huang, D.Y.; Ichikawa, Y.
Biochim. Biophys. Acta 1243, 431-436, 1995
A;Title: Identification of essential lysyl and cysteinyl residues, and the amino acid se
A;Reference number: S60354; MUID:95244596; PMID:7727518
A;Accession: S60354
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <HUA>

Query Match 27.8%; Score 22; DB 2; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGRVH 6
|| ||
Db 1 GGDVH 5

RESULT 8
A60476
S-layer protein - Bacillus thuringiensis (fragment)
C;Species: Bacillus thuringiensis
C;Date: 20-Feb-1993 #sequence_revision 20-Feb-1993 #text_change 09-Jul-2004
C;Accession: A60476
R;Luckevich, M.D.; Beveridge, T.J.

```
J. Bacteriol. 171, 6656-6667, 1989
A;Title: Characterization of a dynamic S layer on Bacillus thuringiensis.
A;Reference number: A60476; MUID:90078111; PMID:2592346
A;Accession: A60476
A;Molecule type: protein
A;Residues: 1-10 <LUC>
A;Cross-references: UNIPROT:P49325
C;Comment: The S-layer, or surface array, is the outermost component of several archaeba

  Query Match             26.6%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 FKDISP 13
              | |: |
Db          5 FPDVXP 10


RESULT 9
PC2330
cycloinulooligosaccharide fructanotransferase (EC 2.4.-.-) - Bacillus circulans (MCI-255
C;Species: Bacillus circulans
C;Date: 21-Mar-1995 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
C;Accession: PC2330
R;Kushibe, S.; Mitsui, K.; Yamagishi, M.; Yamada, K.; Morimoto, Y.
Biosci. Biotechnol. Biochem. 59, 31-34, 1995
A;Title: Purification and characterization of cycloinulooligosaccharide fructanotransfer
A;Reference number: PC2330; MUID:95201377; PMID:7765973
A;Accession: PC2330
A;Molecule type: protein
A;Residues: 1-11 <KUS>
A;Cross-references: UNIPROT:Q7M0L3
C;Comment: This enzyme hydrolyzes beta-(2-1) glycosidic linkages and acts in intermolecu
C;Keywords: glycosyltransferase; hexosyltransferase

  Query Match             26.6%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   37.5%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 HFPKDISP 13
              | |  ::|
Db          2 HLFYQMNP 9


RESULT 10
S01749
collagen alpha 1(I) chain - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S01749; S21415
R;Mooslehner, K.; Harbers, K.
Nucleic Acids Res. 16, 773, 1988
A;Title: Two mRNAs of mouse pro alpha-1(I) collagen gene differ in the size of the 3'-un
A;Reference number: S01749; MUID:88124276; PMID:3340560
A;Accession: S01749
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-12 <MOO>
A;Cross-references: UNIPROT:Q925V7; EMBL:X06753; NID:g50499; PIDN:CAA29927.1; PID:g50500
R;Mooslehner, K.; Harbers, K.
submitted to the EMBL Data Library, July 1989
A;Reference number: S21415
A;Accession: S21415
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-12 <MO2>
A;Cross-references: EMBL:X15896; NID:g50497; PIDN:CAA33904.1; PID:g50498
C;Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;

  Query Match             26.6%;  Score 21;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 FFKDISPI 14
              |  || |:
Db          2 FGLDIGPV 9


RESULT 11
PN0151
omega-gliadine 2' - Aegilops longissima (fragment)
C;Species: Aegilops longissima
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PN0151
R;Odintsova, T.I.; Egorov, T.A.
Biokhimiia 55, 509-516, 1990
A;Title: N-Terminal sequences of omega-gliadins of Aegilops longissima: On the origin of
A;Reference number: PN0146; MUID:90283493; PMID:2354218
A;Accession: PN0151
A;Molecule type: protein
A;Residues: 1-14 <ODI>
A;Experimental source: strain K-907

  Query Match             26.6%;  Score 21;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KDISPI 14
              : ||||
Db          2 RQISPI 7


RESULT 12
PT0670
T-cell receptor beta chain V-D-J region (121-1BN) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0670
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0670
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-9 <FEE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             25.3%;  Score 20;  DB 2;  Length 9;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IGGRV 5
              :|||:
Db          5 LGGRM 9


RESULT 13
S15755
actin 7 - soybean (fragment)
C;Species: Glycine max (soybean)
C;Date: 20-Feb-1995 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: S15755
R;Pearson, L.; Meagher, R.B.
Plant Mol. Biol. 14, 513-526, 1990
A;Title: Diverse soybean actin transcripts contain a large intron in the 5' untranslated
A;Reference number: S15754; MUID:91346640; PMID:2102831
A;Accession: S15755
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-13 <PEA>
A;Cross-references: UNIPROT:P15987; EMBL:X17120; NID:g18527; PIDN:CAA34980.1; PID:g1852
C;Superfamily: actin
C;Keywords: cytoskeleton; structural protein
```

```
  Query Match            25.3%;  Score 20;  DB 2;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 3e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KDISPI 14
              :|| |:
Db          5 EDIQPL 10


RESULT 14
PH0804
T-cell receptor alpha chain (L4) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0804
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0804
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60913
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match            25.3%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   57.1%;  Pred. No. 3.2e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GGRVHFF 8
              |||   |
Db          7 GGRALIF 13


RESULT 15
PC4376
telomeric and tetraplex DNA binding protein qTBP42 VII - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 28-Oct-1997 #sequence_revision 28-Oct-1997 #text_change 09-Jul-2004
C;Accession: PC4376
R;Sarig, G.; Weisman-Shomer, P.; Fry, M.
Biochem. Biophys. Res. Commun. 237, 617-623, 1997
A;Title: Telomeric and tetraplex DNA binding properties of qTBP42: A homologue of the CA
A;Reference number: PC4371; MUID:97445086; PMID:9299414
A;Accession: PC4376
A;Molecule type: protein
A;Residues: 1-14 <SAR>
A;Cross-references: UNIPROT:Q9QX80; UNIPROT:Q9Z0U8; UNIPROT:O88311; UNIPROT:Q9QX81
C;Comment: This protein binds either strand of the telomeric DNA as well as unimolecular

  Query Match            25.3%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 3.2e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          8 FKDISPI 14
              ||:  |:
Db          7 FKEEDPV 13


Search completed: February 22, 2005, 09:46:28
Job time : 11.1333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-11
Perfect score: 79
Sequence: 1 IGGRVHFFKDISPIA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	30.4	11	2	Q70Y70	Q70y70 thorncrofti
2	24	30.4	11	2	Q6YBF2	Q6ybf2 streptococc
3	23.5	29.7	14	2	Q71H50	Q71h50 andrena acc
4	23	29.1	10	1	AKHX_LOCMI	P81626 locusta mig
5	23	29.1	14	2	Q71GY0	Q71gy0 andrena ofe
6	23	29.1	14	2	Q71H20	Q71h20 andrena dis
7	23	29.1	14	2	Q7M0Q6	Q7m0q6 thermotoga
8	23	29.1	15	2	Q9TWF3	Q9twf3 dictyosteli
9	22	27.8	10	1	FARP_MYTED	P42560 mytilus edu
10	22	27.8	12	2	Q88575	Q88575 theiler's e
11	22	27.8	12	2	Q88576	Q88576 theiler's e
12	22	27.8	12	2	Q88577	Q88577 theiler's e
13	22	27.8	12	2	Q88578	Q88578 theiler's e
14	22	27.8	12	2	Q88579	Q88579 theiler's e
15	22	27.8	12	2	Q88580	Q88580 theiler's e
16	22	27.8	12	2	Q88581	Q88581 theiler's e
17	22	27.8	12	2	Q88582	Q88582 theiler's e
18	22	27.8	14	2	Q71GL0	Q71gl0 andrena ili
19	22	27.8	14	2	Q71GS8	Q71gs8 andrena aff
20	22	27.8	14	2	Q71GZ6	Q71gz6 andrena kri
21	22	27.8	14	2	Q71H10	Q71h10 andrena hel
22	22	27.8	14	2	Q71H46	Q71h46 andrena ali
23	22	27.8	15	2	Q9MYT7	Q9myt7 sus scrofa
24	21	26.6	8	2	Q70Y68	Q70y68 prostanther
25	21	26.6	10	1	SLAP_BACTG	P49325 bacillus th
26	21	26.6	10	2	Q9UCU6	Q9ucu6 homo sapien
27	21	26.6	10	2	Q8UVW2	Q8uvw2 oreochromis
28	21	26.6	10	2	Q6R7V4	Q6r7v4 carlia zuma
29	21	26.6	11	2	Q7M0L3	Q7m0l3 bacillus ci
30	21	26.6	11	2	Q68LE0	Q68le0 pyriglena l
31	21	26.6	11	2	Q68LE9	Q68le9 myrmotherul
32	21	26.6	12	2	Q925V7	Q925v7 mus musculu
33	21	26.6	12	2	Q8UVV8	Q8uvv8 squalus aca
34	21	26.6	12	2	Q8UVW0	Q8uvw0 rana catesb
35	21	26.6	14	2	Q6R7V0	Q6r7v0 carlia viva
36	21	26.6	15	2	Q9QV01	Q9qv01 mus sp. . 1
37	20	25.3	10	2	Q9TRS3	Q9trs3 sus scrofa
38	20	25.3	13	1	ACT7_SOYBN	P15987 glycine max
39	20	25.3	13	1	IDHC_PIG	P20304 sus scrofa
40	20	25.3	13	2	Q9L8K1	Q9l8k1 enterococcu
41	20	25.3	13	2	Q8JJ32	Q8jj32 ficedula al
42	20	25.3	13	2	Q00789	Q00789 human t-lym
43	20	25.3	13	2	Q76R60	Q76r60 human t-lym
44	20	25.3	14	1	NEJ2_FASHE	P80526 fasciola he
45	20	25.3	14	2	Q9JHK8	Q9jhk8 mus musculu

ALIGNMENTS

RESULT 1
Q70Y70
ID Q70Y70 PRELIMINARY; PRT; 11 AA.
AC Q70Y70;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Ribosomal protein (Fragment).
GN Name=rps16;
OS Thorncroftia longiflora.
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamiids; Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Thorncroftia.
OX NCBI_TaxID=204202;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15019625; DOI=10.1016/j.ympev.2003.08.002;
RA Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Harley M.M.,
RA Willis F., Simonds M.S.J., Powell M.P., Savolainen V.;
RT "Phylogeny and evolution of basils and allies (Ocimeae, Labiatae)
RT based on three plastid DNA regions.";
RL Mol. Phylogenet. Evol. 31:277-299(2004).
DR EMBL; AJ505401; CAD45521.1; -.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW Ribosomal protein.
FT NON_TER 1 1
FT NON_TER 11 11
SQ SEQUENCE 11 AA; 1201 MW; 396BE78821F2C058 CRC64;

Query Match 30.4%; Score 24; DB 2; Length 11;
Best Local Similarity 57.1%; Pred. No. 3.5e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          2 GGRVHFF 8
              || || :
Db          3 GGIVHLY 9
```

RESULT 2
Q6YBF2
ID Q6YBF2 PRELIMINARY; PRT; 11 AA.
AC Q6YBF2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Wzx (Fragment).
GN Name=wzx;
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.

```
RC   STRAIN=NZSPN00/353, and 01S319;
RX   MEDLINE=22979121; PubMed=14614062; DOI=10.1099/jmm.0.05277-0;
RA   Kong F., Gilbert G.L.;
RT   "Using cpsA-cpsB sequence polymorphisms and serotype-/group-specific
RT   PCR to predict 51 Streptococcus pneumoniae capsular serotypes.";
RL   J. Med. Microbiol. 52:1047-1058(2003).
DR   EMBL; AY163225; AAO60516.1; -.
DR   EMBL; AY163224; AAO60514.1; -.
FT   NON_TER       1       1
SQ   SEQUENCE   11 AA;  1475 MW;  6C6B29E7C721F1B4 CRC64;

  Query Match             30.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 3.5e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 RVHFFK 9
              | | ||
Db          4 RTHLFK 9


RESULT 3
Q71H50
ID   Q71H50      PRELIMINARY;      PRT;    14 AA.
AC   Q71H50;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cytochrome oxidase subunit I (Fragment).
OS   Andrena accepta.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC   Andrenidae; Andreninae; Andrena.
OX   NCBI_TaxID=205209;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Larkin L.L., Neff J.L., Simpson B.B.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF504314; AAQ07599.1; -.
DR   EMBL; AF504313; AAQ07597.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       1       1
SQ   SEQUENCE   14 AA;  1707 MW;  6852BA42A4FF1D5B CRC64;

  Query Match             29.7%;  Score 23.5;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 5.5e+03;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    1;  Gaps    1;

Qy          6 HFFKDISPI 14
              | ||:| ||
Db          3 HSFKEI-PI 10


RESULT 4
AKHX_LOCMI
ID   AKHX_LOCMI     STANDARD;      PRT;    10 AA.
AC   P81626;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Peptide hormone.
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC   Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Corpora cardiaca;
RA   Siegert K.J.;
RL   Submitted (DEC-1998) to Swiss-Prot.
CC   -!- FUNCTION: Probably involved in the regulation of locust
CC       intermediary metabolism, behavior and/or development.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the AKH / HRTH / RPCH family.
DR   InterPro; IPR002047; AKH.
DR   PROSITE; PS00256; AKH; 1.
KW   Amidation; Direct protein sequencing; Neuropeptide;
KW   Pyrrolidone carboxylic acid.
FT   MOD_RES       1       1       Pyrrolidone carboxylic acid.
FT   MOD_RES      10      10       Proline amide.
SQ   SEQUENCE   10 AA;  1222 MW;  81BFF67AB415B9D1 CRC64;

  Query Match             29.1%;  Score 23;  DB 1;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 4.8e+03;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 RVHFFKDISP 13
              :| | :| ||
Db          1 QVTFSRDWSP 10


RESULT 5
Q71GY0
ID   Q71GY0      PRELIMINARY;      PRT;    14 AA.
AC   Q71GY0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cytochrome oxidase subunit I.
OS   Andrena ofella.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC   Andrenidae; Andreninae; Andrena.
OX   NCBI_TaxID=205244;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Larkin L.L., Neff J.L., Simpson B.B.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF504349; AAQ07669.1; -.
DR   EMBL; AF504348; AAQ07667.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
SQ   SEQUENCE   14 AA;  1679 MW;  7463A554622E595B CRC64;

  Query Match             29.1%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 6.8e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 HFFKDISPIA 15
              | | :|  ||
Db          3 HSFNEIPMIA 12


RESULT 6
Q71H20
ID   Q71H20      PRELIMINARY;      PRT;    14 AA.
AC   Q71H20;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Cytochrome oxidase subunit I (Fragment).
OS   Andrena discreta.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC   Andrenidae; Andreninae; Andrena.
OX   NCBI_TaxID=205225;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Larkin L.L., Neff J.L., Simpson B.B.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF504329; AAQ07629.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
```

```
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1622 MW;  747525547677BE5B CRC64;

  Query Match             29.1%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 6.8e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          6 HFFKDISPIA 15
              | | :|  ||
Db          3 HSFNEIPVIA 12


RESULT 7
Q7M0Q6
ID   Q7M0Q6      PRELIMINARY;      PRT;    14 AA.
AC   Q7M0Q6;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Xylan 1,4-beta-xylosidase (EC 3.2.1.37) (Fragment).
OS   Thermotoga sp.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=28240;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93152594; PubMed=8427876; DOI=10.1016/0304-4165(93)90132-R;
RA   Ruttersmith L.D., Daniel R.M.;
RT   "Thermostable beta-glucosidase and beta-xylosidase from Thermotoga sp.
RT   strain FjSS3-B.1.";
RL   Biochim. Biophys. Acta 1156:167-172(1993).
DR   PIR; S29632; S29632.
DR   GO; GO:0009044; F:xylan 1,4-beta-xylosidase activity; IEA.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1709 MW;  CB22131EE55ADCC9 CRC64;

  Query Match             29.1%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 6.8e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 VHFFKD 10
              |:|| |
Db          5 VYFFAD 10


RESULT 8
Q9TWF3
ID   Q9TWF3      PRELIMINARY;      PRT;    15 AA.
AC   Q9TWF3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   V-ATPase D-subunit (Fragment).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=96019266; PubMed=7593293;
RA   Adessi C., Chapel A., Vincon M., Rabilloud T., Klein G., Satre M.,
RA   Garin J.;
RT   "Identification of major proteins associated with Dictyostelium
RT   discoideum endocytic vesicles.";
RL   J. Cell Sci. 108:3331-3337(1995).
SQ   SEQUENCE   15 AA;  1608 MW;  E4A91C650AA853EA CRC64;

  Query Match             29.1%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 7.2e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GGRVH 6
              ||| |
Db          4 GGRKH 8


RESULT 9
FARP_MYTED
ID   FARP_MYTED     STANDARD;      PRT;    10 AA.
AC   P42560;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide ALAGDHFFRF-amide.
OS   Mytilus edulis (Blue mussel).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;
OC   Mytiloidea; Mytilidae; Mytilus.
OX   NCBI_TaxID=6550;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=93047883; PubMed=1358534; DOI=10.1016/0742-8413(92)90104-F;
RA   Walker R.J.;
RT   "Neuroactive peptides with an RFamide or Famide carboxyl terminal.";
RL   Comp. Biochem. Physiol. 102C:213-222(1992).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
DR   PIR; A58365; A58365.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES      10     10       Phenylalanine amide.
SQ   SEQUENCE   10 AA;  1180 MW;  C2F80CC9C1EAA87D CRC64;

  Query Match             27.8%;  Score 22;  DB 1;  Length 10;
  Best Local Similarity   75.0%;  Pred. No. 7.2e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 HFFK 9
              |||:
Db          6 HFFR 9


RESULT 10
Q88575
ID   Q88575      PRELIMINARY;      PRT;    12 AA.
AC   Q88575;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Theiler's murine encephalomyelitis virus 5' flank and 5' end cds.
DE   (Fragment).
OS   Theiler's encephalomyelitis virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Cardiovirus.
OX   NCBI_TaxID=12124;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FA;
RX   MEDLINE=92194426; PubMed=1548749;
RA   Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;
RT   "A single base deletion in the 5' noncoding region of Theiler's virus
RT   attenuates neurovirulence.";
RL   J. Virol. 66:1951-1958(1992).
DR   EMBL; M80883; AAA73154.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1321 MW;  F79B7D908E7B5871 CRC64;

  Query Match             27.8%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 8.7e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 HFFKDISP 13
              | : |: |
Db          5 HGYPDVCP 12


RESULT 11
```

```
Q88576
ID   Q88576      PRELIMINARY;      PRT;    12 AA.
AC   Q88576;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Theiler's murine encephalomyelitis virus 5' flank and 5' end cds.
DE   (Fragment).
OS   Theiler's encephalomyelitis virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Cardiovirus.
OX   NCBI_TaxID=12124;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MHG;
RX   MEDLINE=92194426; PubMed=1548749;
RA   Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;
RT   "A single base deletion in the 5' noncoding region of Theiler's virus
RT   attenuates neurovirulence.";
RL   J. Virol. 66:1951-1958(1992).
DR   EMBL; M80884; AAA73155.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1321 MW;  F79B7D908E7B5871 CRC64;

  Query Match             27.8%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 8.7e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 HFFKDISP 13
              | : |: |
Db          5 HGYPDVCP 12


RESULT 12
Q88577
ID   Q88577      PRELIMINARY;      PRT;    12 AA.
AC   Q88577;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Theiler's murine encephalomyelitis virus 5' flank and 5' end cds.
DE   (Fragment).
OS   Theiler's encephalomyelitis virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Cardiovirus.
OX   NCBI_TaxID=12124;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TO;
RX   MEDLINE=92194426; PubMed=1548749;
RA   Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;
RT   "A single base deletion in the 5' noncoding region of Theiler's virus
RT   attenuates neurovirulence.";
RL   J. Virol. 66:1951-1958(1992).
DR   EMBL; M80885; AAA73156.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1321 MW;  F79B7D908E7B5871 CRC64;

  Query Match             27.8%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 8.7e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 HFFKDISP 13
              | : |: |
Db          5 HGYPDVCP 12


RESULT 13
Q88578
ID   Q88578      PRELIMINARY;      PRT;    12 AA.
AC   Q88578;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Theiler's murine encephalomyelitis virus 5' flank and 5' end cds.
DE   (Fragment).
OS   Theiler's encephalomyelitis virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Cardiovirus.
OX   NCBI_TaxID=12124;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TO;
RX   MEDLINE=92194426; PubMed=1548749;
RA   Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;
RT   "A single base deletion in the 5' noncoding region of Theiler's virus
RT   attenuates neurovirulence.";
RL   J. Virol. 66:1951-1958(1992).
DR   EMBL; M80886; AAA73157.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1321 MW;  F79B7D908E7B5871 CRC64;

  Query Match             27.8%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 8.7e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 HFFKDISP 13
              | : |: |
Db          5 HGYPDVCP 12


RESULT 14
Q88579
ID   Q88579      PRELIMINARY;      PRT;    12 AA.
AC   Q88579;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Theiler's murine encephalomyelitis virus 5' flank and 5' end cds.
DE   (Fragment).
OS   Theiler's encephalomyelitis virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Cardiovirus.
OX   NCBI_TaxID=12124;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VL;
RX   MEDLINE=92194426; PubMed=1548749;
RA   Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;
RT   "A single base deletion in the 5' noncoding region of Theiler's virus
RT   attenuates neurovirulence.";
RL   J. Virol. 66:1951-1958(1992).
DR   EMBL; M80887; AAA73158.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1321 MW;  F79B7D908E7B5871 CRC64;

  Query Match             27.8%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 8.7e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 HFFKDISP 13
              | : |: |
Db          5 HGYPDVCP 12


RESULT 15
Q88580
ID   Q88580      PRELIMINARY;      PRT;    12 AA.
AC   Q88580;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Theiler's murine encephalomyelitis virus 5' flank and 5' end cds.
DE   (Fragment).
OS   Theiler's encephalomyelitis virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
```

```
OC   Cardiovirus.
OX   NCBI_TaxID=12124;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Vilyuisk;
RX   MEDLINE=92194426; PubMed=1548749;
RA   Pritchard A.E., Calenoff M.A, Simpson S., Jensen K., Lipton H.L.;
RT   "A single base deletion in the 5' noncoding region of Theiler's virus
RT   attenuates neurovirulence.";
RL   J. Virol. 66:1951-1958(1992).
DR   EMBL; M80888; AAA73159.1; -.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1321 MW;  F79B7D908E7B5871 CRC64;

  Query Match             27.8%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   37.5%;  Pred. No. 8.7e+03;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 HFFKDISP 13
              | : |: |
Db          5 HGYPDVCP 12


Search completed: February 22, 2005, 09:38:01
Job time : 54.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-11
Perfect score: 79
Sequence: 1 IGGRVHFFKDISPIA 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	79	100.0	15	2	AAW04851	Aaw04851 Internal
2	39	49.4	13	7	ADC51728	Adc51728 Human MBP
3	39	49.4	13	7	ADC51729	Adc51729 Human MBP
4	37	46.8	11	4	AAB70569	Aab70569 Human imm
5	37	46.8	11	5	AAO21025	Aao21025 Isomerise
6	37	46.8	12	3	AAB26883	Aab26883 Human mye
7	35	44.3	13	2	AAR74902	Aar74902 Myelin Ba
8	35	44.3	13	5	ABB84347	Abb84347 MBP epito
9	35	44.3	15	3	AAY85541	Aay85541 Human MBP
10	35	44.3	15	3	AAB12643	Aab12643 Modified
11	35	44.3	15	4	AAM99034	Aam99034 Vaccine r
12	35	44.3	15	4	AAM99033	Aam99033 Vaccine r
13	34	43.0	7	2	AAW72354	Aaw72354 Human mye
14	34	43.0	8	2	AAR61755	Aar61755 MBP pepti
15	34	43.0	8	8	ADR48840	Adr48840 Inhibitor
16	34	43.0	8	8	ADR48797	Adr48797 Inhibitor
17	34	43.0	9	2	AAR61856	Aar61856 MBP pepti
18	34	43.0	9	2	AAR61871	Aar61871 MBP pepti
19	34	43.0	9	2	AAR91918	Aar91918 Generic p
20	34	43.0	9	8	ADM12852	Adm12852 MHC class
21	34	43.0	9	8	ADO39090	Ado39090 Human mye
22	34	43.0	10	2	AAR61949	Aar61949 MBP pepti
23	34	43.0	10	2	AAR61952	Aar61952 MBP pepti
24	34	43.0	10	2	AAR91921	Aar91921 Peptide c
25	34	43.0	10	2	AAR91919	Aar91919 Peptide c
26	34	43.0	10	2	AAR91920	Aar91920 Peptide c
27	34	43.0	10	2	AAR91922	Aar91922 Peptide c
28	34	43.0	10	2	AAW72358	Aaw72358 Human mye
29	34	43.0	10	2	AAW72355	Aaw72355 Human mye
30	34	43.0	10	2	AAW72357	Aaw72357 Human mye
31	34	43.0	10	2	AAW72356	Aaw72356 Human mye
32	34	43.0	10	7	ADB89021	Adb89021 Human mye
33	34	43.0	10	7	ADC06706	Adc06706 Myelin ba
34	34	43.0	10	7	ADE38270	Ade38270 Immunomod
35	34	43.0	11	2	AAW54673	Aaw54673 Peptide f
36	34	43.0	12	8	ADN65365	Adn65365 HLA bindi
37	34	43.0	13	2	AAR29955	Aar29955 TCR pepti
38	34	43.0	13	2	AAR74910	Aar74910 Myelin Ba
39	34	43.0	13	2	AAR74911	Aar74911 Myelin Ba
40	34	43.0	13	2	AAR74909	Aar74909 Myelin Ba
41	34	43.0	13	2	AAR74904	Aar74904 Myelin Ba
42	34	43.0	13	2	AAR74903	Aar74903 Myelin Ba
43	34	43.0	13	2	AAR85133	Aar85133 Human MBP
44	34	43.0	13	2	AAY04452	Aay04452 Myelin ba
45	34	43.0	13	2	AAR95363	Aar95363 Residues

ALIGNMENTS

RESULT 1
AAW04851
ID AAW04851 standard; peptide; 15 AA.
XX
AC AAW04851;
XX
DT 27-AUG-2003 (revised)
DT 18-FEB-1997 (first entry)
XX
DE Internal fragment of human papilloma virus type 7 L2 protein.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Human papillomavirus type 7.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 2; Page 44; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myelin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with

```
CC   the autoimmune disease. This peptide is an internal peptide of human
CC   papilloma virus type 7 L2 protein and is implicated as a foreign epitope
CC   involved in the aetiology or in remissions of multiple sclerosis. It has
CC   been shown capable of inducing the proliferation of autoreactive T-cell
CC   clones isolated from multiple sclerosis patients. (Updated on 27-AUG-2003
CC   to correct OS field.)
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 79;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 1e-07;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 IGGRVHFFKDISPIA 15
              |||||||||||||||
Db          1 IGGRVHFFKDISPIA 15


RESULT 2
ADC51728
ID   ADC51728 standard; peptide; 13 AA.
XX
AC   ADC51728;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human MBP87-99 peptide #2.
XX
KW   human; PCMT1; IAMT; L-Isoaspartyl (D-asparty1)O-methyltransferase; IAMT;
KW   autoimmune; celiac disease; Crohn's disease;
KW   insulin dependent diabetes mellitus; Grave's disease; multiple sclerosis;
KW   myasthenia gravis; psoriasis; rheumatoid arthritis; Sjogren's syndrome;
KW   systemic lupus erythematosus; ulcerative colitis; MBP87-99.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   6
FT                   /label= Diso
XX
PN   WO2003057204-A2.
XX
PD   17-JUL-2003.
XX
PF   07-JAN-2003; 2003WO-EP000079.
XX
PR   08-JAN-2002; 2002DK-00000026.
PR   08-JAN-2002; 2002US-0346709P.
XX
PA   (NORD-) NORDIC BIOSCIENCE AS.
XX
PI   Cloos PAC;
XX
DR   WPI; 2003-645976/61.
XX
PT   Use of L-Isoaspartyl (D-asparty1)O-methyltransferase (IAMT) activity
PT   regulator, IAMT, IAMT encoding nucleic acid sequence or their functional
PT   derivatives for preparing composition for the treatment of an autoimmune
PT   response and/or disease.
XX
PS   Example 1; Fig 3; 34pp; English.
XX
CC   The invention relates to the novel use of a regulator of L-Isoaspartyl (D
CC   -asparty1)O-methyltransferase (IAMT) activity (A) or IAMT or IAMT
CC   encoding nucleic acid sequence or their functional derivatives in the
CC   preparation of a composition for the prevention, treatment or alleviation
CC   of an autoimmune response and/or disease. The composition of the
CC   invention is used for the prevention, treatment or alleviation of an
CC   autoimmune response and/or disease in a mammal e.g. celiac disease,
CC   Crohn's disease, insulin dependent diabetes mellitus, Grave's disease,
CC   multiple sclerosis, myasthenia gravis, psoriasis, rheumatoid arthritis,
CC   Sjogren's syndrome, systemic lupus erythematosus or ulcerative colitis.
CC   The present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 13 AA;

  Query Match             49.4%;  Score 39;  DB 7;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VHFFKDI 11
              |||||||
Db          1 VHFFKDI 7


RESULT 3
ADC51729
ID   ADC51729 standard; peptide; 13 AA.
XX
AC   ADC51729;
XX
DT   18-DEC-2003  (first entry)
XX
DE   Human MBP87-99 peptide #3.
XX
KW   human; PCMT1; IAMT; L-Isoaspartyl (D-asparty1)O-methyltransferase; IAMT;
KW   autoimmune; celiac disease; Crohn's disease;
KW   insulin dependent diabetes mellitus; Grave's disease; multiple sclerosis;
KW   myasthenia gravis; psoriasis; rheumatoid arthritis; Sjogren's syndrome;
KW   systemic lupus erythematosus; ulcerative colitis; MBP87-99.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   6
FT                   /note= "D-form residue"
XX
PN   WO2003057204-A2.
XX
PD   17-JUL-2003.
XX
PF   07-JAN-2003; 2003WO-EP000079.
XX
PR   08-JAN-2002; 2002DK-00000026.
PR   08-JAN-2002; 2002US-0346709P.
XX
PA   (NORD-) NORDIC BIOSCIENCE AS.
XX
PI   Cloos PAC;
XX
DR   WPI; 2003-645976/61.
XX
PT   Use of L-Isoaspartyl (D-asparty1)O-methyltransferase (IAMT) activity
PT   regulator, IAMT, IAMT encoding nucleic acid sequence or their functional
PT   derivatives for preparing composition for the treatment of an autoimmune
PT   response and/or disease.
XX
PS   Example 1; Fig 3; 34pp; English.
XX
CC   The invention relates to the novel use of a regulator of L-Isoaspartyl (D
CC   -asparty1)O-methyltransferase (IAMT) activity (A) or IAMT or IAMT
CC   encoding nucleic acid sequence or their functional derivatives in the
CC   preparation of a composition for the prevention, treatment or alleviation
CC   of an autoimmune response and/or disease. The composition of the
CC   invention is used for the prevention, treatment or alleviation of an
CC   autoimmune response and/or disease in a mammal e.g. celiac disease,
CC   Crohn's disease, insulin dependent diabetes mellitus, Grave's disease,
CC   multiple sclerosis, myasthenia gravis, psoriasis, rheumatoid arthritis,
CC   Sjogren's syndrome, systemic lupus erythematosus or ulcerative colitis.
CC   The present sequence is used in the exemplification of the invention.
XX
SQ   Sequence 13 AA;

  Query Match             49.4%;  Score 39;  DB 7;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 3.3;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          5 VHPFKDI 11
              |||||||
Db          1 VHPFKDI 7
```

RESULT 4
AAB70569
ID AAB70569 standard; peptide; 11 AA.
XX
AC AAB70569;
XX
DT 10-MAY-2001 (first entry)
XX
DE Human immunoglobulin G (IgG) epitope #10.
XX
KW Immunoglobulin G; IgG; myelin basic protein; MBP; MOG; Crohn's disease;
KW myelin oligodendrocyte glycoprotein; alphaB-crystallin; MS; psoriasis;
KW multiple sclerosis; rheumatoid arthritis; autoimmune disease; IDDM;
KW immune system; insulin dependent diabetes mellitus; myasthenia gravis;
KW coeliac disease; Chaga's disease; diagnosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 7
FT /note= "Asx can be alpha-D Asn or Asp, or is beta-D or
FT beta-L Asp"
XX
PN WO200113110-A2.
XX
PD 22-FEB-2001.
XX
PF 16-AUG-2000; 2000WO-EP007973.
XX
PR 17-AUG-1999; 99GB-00019452.
XX
PA (OSTE-) OSTEOMETER BIOTECH AS.
XX
PI Cloos PAC, Christgau S;
XX
DR WPI; 2001-234923/24.
XX
PT Diagnostic assay useful for quantitative/qualitative determination of
PT autoimmune reactivity comprises determining an immune system component
PT recognizing epitope containing isomerized peptide linkage and/or
PT optically inverted amino acid.
XX
PS Claim 7; Page 64; 74pp; English.
XX
CC The present invention describes a diagnostic assay (M1) comprising
CC quantitative or qualitative determination of an auto-reactive immune
CC system component specifically recognising an epitope (E) containing an
CC isomerised peptide linkage and/or optically inverted amino acid, an auto-
CC antigen or its fragment containing (E) and/or a non-self antigen or its
CC fragment which contains (E) and is capable of inducing an autoimmune
CC response. The diagnostic assay is useful for quantitative and qualitative
CC determination of autoimmune reactivity of an isomerised protein sequence,
CC which is an indicative of an autoimmune disease such as rheumatoid
CC arthritis, multiple sclerosis, insulin dependent diabetes mellitus,
CC myasthenia gravis, coeliac disease, Chaga's disease, psoriasis or Crohn's
CC disease. The present sequence represents a specifically claimed example
CC of (E) for use in the diagnostic assay
XX
SQ Sequence 11 AA;

Query Match 46.8%; Score 37; DB 4; Length 11;
Best Local Similarity 85.7%; Pred. No. 6.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          5 VHPFKDI 11
              |||||:|
Db          2 VHPFKBI 8
```

RESULT 5
AAO21025
ID AAO21025 standard; peptide; 11 AA.
XX
AC AAO21025;
XX
DT 19-JUL-2002 (first entry)
XX
DE Isomerised/optically inverted epitope #8 for recognition by antibody.
XX
KW Immunosuppressive; antirheumatic; antiarthritic; antidiabetic; epitope;
KW antithyroid; neuroprotective; antiinflammatory; autoimmune disease;
KW antipsoriatic; selective suppression; auto-antigen; immunological;
KW isomerised peptide linkage; optionally inverted; rheumatoid arthritis;
KW vaccine; multiple sclerosis; insulin dependent diabetes mellitus;
KW myasthenia gravis; celiac disease; Chagas' disease; psoriasis;
KW Crohn's disease.
XX
OS Unidentified.
XX
PN WO200213844-A2.
XX
PD 21-FEB-2002.
XX
PF 09-AUG-2001; 2001WO-EP009205.
XX
PR 16-AUG-2000; 2000GB-00020238.
XX
PA (OSTE-) OSTEOMETER BIOTECH AS.
XX
PI Cloos PAC, Christgau S;
XX
DR WPI; 2002-339411/37.
XX
PT Use of protein, peptide or their analogs containing an epitope recognized
PT by an auto-reactive immune system for the treatment of autoimmune
PT disease.
XX
PS Claim 5; Page 92; 111pp; English.
XX
CC The invention relates to methods and compositions for the treatment of an
CC autoimmune disease involving the selective suppression of autoimmune
CC activity against an auto-antigen. The activity involves immunological
CC reactivity with an epitope containing an isomerised peptide linkage and/
CC or an optionally inverted amino acid. The epitopes of the invention can
CC be used for therapeutic or prophylactic treatment of the autoimmune
CC diseases for the vaccination against autoimmune diseases e.g. rheumatoid
CC arthritis, multiple sclerosis, insulin dependent diabetes mellitus,
CC myasthenia gravis, celiac disease, Chagas' diseases, psoriasis and
CC Crohn's disease. This sequence represents an isomerised/optically
CC inverted epitope of the invention
XX
SQ Sequence 11 AA;

Query Match 46.8%; Score 37; DB 5; Length 11;
Best Local Similarity 85.7%; Pred. No. 6.7;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          5 VHPFKDI 11
              |||||:|
Db          2 VHPFKBI 8
```

RESULT 6
AAB26883
ID AAB26883 standard; peptide; 12 AA.
XX
AC AAB26883;
XX
DT 01-FEB-2001 (first entry)
XX

```
DE   Human myelin basic protein fragment.
XX
KW   Immunosuppressant; neuroprotective; multiple sclerosis; treatment.
XX
OS   Homo sapiens.
XX
PN   WO200058354-A1.
XX
PD   05-OCT-2000.
XX
PF   29-MAR-1999;   99WO-EP002268.
XX
PR   29-MAR-1999;   99WO-EP002268.
XX
PA   (TECN-) TECNOGEN SCPA.
XX
PI   Marino M,  Ippolito A,  Fassina G;
XX
DR   WPI; 2000-628344/60.
XX
PT   New glycine-rich peptides, useful for treatment of multiple sclerosis by
PT   induction of anergy in autoreactive T cells.
XX
PS   Disclosure; Page 2-3; 23pp; English.
XX
CC   Peptides AAB26878-B26881 represent four glycine rich peptide which may be
CC   N-acetylated and/or C-amidated and contain amino acids with L or D
CC   configuration. Included in the invention is a pharmaceutical composition
CC   containing at least one of the peptides and an inert ingredient. The
CC   peptides have immunosuppressant and neuroprotective activity and induce
CC   anergy of autoreactive T lymphocytes without inducing an autoreactive
CC   response. The peptides are used to treat multiple sclerosis. The present
CC   sequence represents a fragment of the human myelin basic protein which is
CC   used in the invention in the design of the peptides of the invention
XX
SQ   Sequence 12 AA;

  Query Match             46.8%;  Score 37;  DB 3;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 7.3;
  Matches    6;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 VHFFKDISP 13
              ||||| ::|
Db          1 VHFFKIVTP 9


RESULT 7
AAR74902
ID   AAR74902 standard; peptide; 13 AA.
XX
AC   AAR74902;
XX
DT   25-MAR-2003  (revised)
DT   29-NOV-1995  (first entry)
XX
DE   Myelin Basic protein epitope alanine substituted peptide ala8.
XX
KW   T-cell Receptor; TcR; variable region; multiple sclerosis;
KW   autoimmune disease; neurodegeneration; myelin basic protein; MBP;
KW   oligopeptide immunogen; competitor peptide; antagonist.
XX
OS   Synthetic.
XX
PN   WO9508572-A1.
XX
PD   30-MAR-1995.
XX
PF   22-SEP-1994;   94WO-US010728.
XX
PR   22-SEP-1993;   93US-00125407.
XX
PA   (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI   Steinman L,  Oksenberg J,  Bernard C,  Zamvil S,  Mitchell DJ;
PI   Karin N;
XX
DR   WPI; 1995-139558/18.
XX
PT   Determining relation between auto:immune degenerative diseases and
PT   specific variable regions of T-cell receptors - as associated with the
PT   host HLA or T-cells associated with combating neoproliferative diseases.
XX
PS   Claim 6; Page 116; 122pp; English.
XX
CC   A set of substituted peptides based on the sequence corresp. to the I-E
CC   restricted epitope MBP 89-99 in the rat was synthesised and tested for
CC   MHC binding. Six of the peptides (see AAR74899-R74904) have substantially
CC   reduced MHC binding affinity, a reduced ability to stimulate T-cells in
CC   vitro and a reduced ability to induce autoimmune disease. They are useful
CC   as competitor peptides to antagonise the T-cell receptor, e.g. to block
CC   the CDR3 region involved in multiple sclerosis. Oligopeptide immunogens
CC   comprising one of the 6 amino acid sequences are also claimed. (Updated
CC   on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 13 AA;

  Query Match             44.3%;  Score 35;  DB 2;  Length 13;
  Best Local Similarity   75.0%;  Pred. No. 19;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VHFFKDIS 12
              |||||:|:
Db          1 VHFFKNIA 8


RESULT 8
ABB84347
ID   ABB84347 standard; peptide; 13 AA.
XX
AC   ABB84347;
XX
DT   17-OCT-2002  (first entry)
XX
DE   MBP epitope analogue SEQ ID 47.
XX
KW   MBP; myelin basic protein; tolerance; immune system; multiple sclerosis;
KW   autoimmune response; autoimmune disease; immunosuppressive;
KW   neuroprotective.
XX
OS   Synthetic.
XX
PN   US2002076412-A1.
XX
PD   20-JUN-2002.
XX
PF   07-JUN-1995;   95US-00484409.
XX
PR   17-AUG-1987;   87US-00086694.
PR   12-JUL-1989;   89US-00379500.
PR   01-MAY-1990;   90US-00517245.
PR   01-MAY-1991;   91WO-US002991.
PR   30-APR-1992;   92US-00877444.
PR   21-MAY-1993;   93US-00066325.
PR   22-SEP-1993;   93US-00125407.
XX
PA   (STEI/) STEINMAN L.
PA   (ZAMV/) ZAMVIL S.
XX
PI   Steinman L,  Zamvil S;
XX
DR   WPI; 2002-598709/64.
XX
PT   Modulating or tolerizing the immune system, useful for treating multiple
PT   sclerosis, by administering a peptide derived from human myelin binding
PT   protein.
XX
```

PS Example II; Page 17; 21pp; English.
XX
CC This invention describes a novel method for modulating or tolerizing the
CC immune system, and for treating multiple sclerosis comprising
CC administering a peptide derived from hMBP (human myelin basic protein).
CC The peptide induces an autoimmune response (T cell) to a self-antigen (or
CC part of it), and binds to an MHC (major histocompatibility complex)
CC antigen of a host susceptible to autoimmune diseases, i.e. competes with
CC binding to MBP and inhibit proliferation of MBP-reactive cells. The
CC peptide has immunosuppressive and neuroprotective activity. This sequence
CC represents a synthetic analogue of the human MBP protein I-E restricted
CC epitope
XX
SQ Sequence 13 AA;

Query Match 44.3%; Score 35; DB 5; Length 13;
Best Local Similarity 75.0%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VHPFKDIS 12
||||||:|:
Db 1 VHPFKNIA 8

RESULT 9
AAY85541
ID AAY85541 standard; peptide; 15 AA.
XX
AC AAY85541;
XX
DT 23-JUN-2000 (first entry)
XX
DE Human MBP peptide variant 96A.
XX
KW MBP; CD4+ T cell; myelin basic protein; multiple sclerosis; HLA;
KW human leukocyte antigen; epitope; cytokine; interleukin-4; variant;
KW transforming growth factorbeta.
XX
OS Homo sapiens.
XX
PN US6036957-A.
XX
PD 14-MAR-2000.
XX
PF 06-JUN-1995; 95US-00469640.
XX
PR 30-MAR-1990; 90US-00502559.
PR 28-FEB-1992; 92US-00843752.
PR 09-APR-1992; 92US-00865318.
PR 09-APR-1993; 93US-00046354.
XX
PA (AUTO-) AUTOIMMUNE INC.
XX
PI Hafler DA, Miller A, Al-Sabbagh A, Weiner HL;
XX
DR WPI; 2000-246181/21.
XX
PT Suppressing T cell response to myelin basic protein, useful for treating
PT multiple sclerosis, by administering a peptide containing immunodominant
PT epitope of the protein.
XX
PS Disclosure; Fig 11; 40pp; English.
XX
CC The invention relates to a method of suppressing the immune function of
CC CD4+ T cells, reactive with myelin basic protein (MBP) in patients with
CC multiple sclerosis. The method comprises administering a peptide (I) that
CC stimulates (in a proliferation assay) the subgroup of HLA (human
CC leukocyte antigen)-DR2b-restricted T clones, from remitting-relapsing
CC patients, that is reactive with another peptide. (I) contains
CC immunodominant epitopes of MBP (residues 84-102) and when administered
CC orally actively suppress T cells reactive with this epitope (by inducing
CC CD8+ suppressor cells that express cytokines such as transforming growth
CC factorbeta and interleukin-4), or when administered intravenously they
CC induce clonal anergy. The method is used for treatment of multiple
CC sclerosis. Sequences AAY85530-545 represent variants of human MBP,
CC derived from residues 85-99
XX
SQ Sequence 15 AA;

Query Match 44.3%; Score 35; DB 3; Length 15;
Best Local Similarity 75.0%; Pred. No. 23;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VHPFKDIS 12
||||||:|:
Db 5 VHPFKNIA 12

RESULT 10
AAB12643
ID AAB12643 standard; peptide; 15 AA.
XX
AC AAB12643;
XX
DT 10-NOV-2000 (first entry)
XX
DE Modified human myelin basic protein peptide #12.
XX
KW Human; myelin basic protein; MBP; multiple sclerosis; MS; CNS;
KW immunodominant; chronic inflammatory disease; central nervous system;
KW autoimmune disease; neuroprotective; immunosuppressive; CD8+; CD4+.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN US6077509-A.
XX
PD 20-JUN-2000.
XX
PF 06-JUN-1995; 95US-00469648.
XX
PR 30-MAR-1990; 90US-00502559.
PR 28-FEB-1992; 92US-00843752.
PR 09-APR-1992; 92US-00865318.
PR 09-APR-1993; 93US-00046354.
XX
PA (AUTO-) AUTOIMMUNE INC.
XX
PI Hafler DA, Weiner HL;
XX
DR WPI; 2000-450923/39.
XX
PT New pharmaceutical composition comprising peptide fragments of myelin
PT basic protein, useful in suppressing the symptoms of multiple sclerosis.
XX
PS Example 4; Fig 11; 26pp; English.
XX
CC The present invention describes a pharmaceutical formulation comprising a
CC peptide having a 30-amino acid sequence (AAB12612) (I) or a segment of
CC (I), provided that the segment comprises at least the 19-amino acid
CC sequence given in AAB12613. The pharmaceutical formulation has
CC neuroprotective and immunosuppressive activities. The compositions and
CC peptides from the present invention are useful in suppressing multiple
CC sclerosis, which is due to the induction of CD8+ suppressor T-cells. The
CC peptides are also useful in identifying CD4+ T-cells reactive with myelin
CC basic protein (MBP). The peptides are also useful in identifying
CC individuals with T-cells reactive to MBP. Oral administration of MBP
CC suppressed central nervous system (CNS) inflammation in actively induced
CC experimental allergic encephalomyelitis (EAE). There was a decreased
CC inflammation in both the parenchyma and meninges when cells from MBP-fed
CC animals were transferred. This suppression was observed when CD4+
CC depleted, but not CD8+. The present sequence represents a modified human
CC MBP peptide fragment, which is used in an example from the present
CC invention
XX
SQ Sequence 15 AA;

```
  Query Match             44.3%;  Score 35;  DB 3;  Length 15;
  Best Local Similarity   75.0%;  Pred. No. 23;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VHFFKDIS 12
              |||||:|:
Db          5 VHFFKNIA 12
```

```
RESULT 11
AAM99034
ID   AAM99034 standard; peptide; 15 AA.
XX
AC   AAM99034;
XX
DT   07-DEC-2001  (first entry)
XX
DE   Vaccine related MHC ligand peptide SEQ ID NO:137.
XX
KW   Glutamic acid; glutamine; vaccine; major histocompatibility complex; MHC;
KW   immunomodulator; antiallergic; endocrine; neuroprotectant; virucidal;
KW   bactericidal; antiparasitic; fungicidal; cytostatic; medicine;
KW   pharmaceutical; immune disorder; immune deficiency; autoimmune;
KW   hypersensitivity; allergy; graft rejection; infection; hormonal disorder;
KW   central nervous system disease; cancer; melanoma; anti-melanoma vaccine;
KW   human immunodeficiency virus.
XX
OS   Homo sapiens.
XX
PN   WO200170772-A2.
XX
PD   27-SEP-2001.
XX
PF   22-MAR-2001; 2001WO-FR000872.
XX
PR   23-MAR-2000; 2000FR-00003711.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI   Klinguer-Hamour C,  Corvaia N,  Beck A,  Goetsch L;
XX
DR   WPI; 2001-611470/70.
XX
PT   Stabilized pharmaceutical containing N-terminal glutamic acid or
PT   glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
PT   with strong acid.
XX
PS   Claim 9; Page 54; 149pp; French.
XX
CC   The present invention describes a pharmaceutical compound (I) that
CC   contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue in
CC   the form of an addition salt with a strong, physiologically acceptable
CC   acid (II). Also described are: (a) a pharmaceutical composition
CC   containing at least one (I); (b) a vaccine containing at least one (I)
CC   where this is a major histocompatibility complex (MHC) ligand (Ia); (c) a
CC   method for in vitro diagnosis of diseases associated with the presence of
CC   (Ia); (d) a kit for method (c) that includes a (Ia); and (e) a process
CC   for preparing (I). (I) has immunomodulator, endocrine, antiallergic,
CC   neuroprotectant, virucidal, bactericidal, antiparasitic, fungicidal and
CC   cytostatic activities. (I) are useful, in human or veterinary medicine,
CC   in pharmaceutical compositions (for treating immune disorders, e.g.
CC   immune deficiency, autoimmune states, hypersensitivity, allergy, graft
CC   rejection, infection, hormonal disorders and central nervous system
CC   diseases), also, where (I) is a MHC ligand (Ia), in vaccines for
CC   treatment or prevention of: (i) viral, bacterial, parasitic or fungal
CC   infections; or (ii) of cancers. A particular application is in anti-
CC   melanoma vaccines. (I) are also useful for in vitro diagnosis of diseases
CC   associated with interactions between MHC and (I), e.g. melanoma and human
CC   immunodeficiency virus infection. AAM98898 to AAM99592 represent peptides
CC   which can be used in pharmaceutical compounds from the present invention
XX
SQ   Sequence 15 AA;
```

```
  Query Match             44.3%;  Score 35;  DB 4;  Length 15;
  Best Local Similarity   75.0%;  Pred. No. 23;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VHFFKDIS 12
              |||||:|:
Db          5 VHFFKNIA 12
```

```
RESULT 12
AAM99033
ID   AAM99033 standard; peptide; 15 AA.
XX
AC   AAM99033;
XX
DT   07-DEC-2001  (first entry)
XX
DE   Vaccine related MHC ligand peptide SEQ ID NO:136.
XX
KW   Glutamic acid; glutamine; vaccine; major histocompatibility complex; MHC;
KW   immunomodulator; antiallergic; endocrine; neuroprotectant; virucidal;
KW   bactericidal; antiparasitic; fungicidal; cytostatic; medicine;
KW   pharmaceutical; immune disorder; immune deficiency; autoimmune;
KW   hypersensitivity; allergy; graft rejection; infection; hormonal disorder;
KW   central nervous system disease; cancer; melanoma; anti-melanoma vaccine;
KW   human immunodeficiency virus.
XX
OS   Homo sapiens.
XX
PN   WO200170772-A2.
XX
PD   27-SEP-2001.
XX
PF   22-MAR-2001; 2001WO-FR000872.
XX
PR   23-MAR-2000; 2000FR-00003711.
XX
PA   (FABR ) FABRE MEDICAMENT SA PIERRE.
XX
PI   Klinguer-Hamour C,  Corvaia N,  Beck A,  Goetsch L;
XX
DR   WPI; 2001-611470/70.
XX
PT   Stabilized pharmaceutical containing N-terminal glutamic acid or
PT   glutamine, useful e.g. in anti-melanoma vaccines, is an addition salt
PT   with strong acid.
XX
PS   Claim 9; Page 54; 149pp; French.
XX
CC   The present invention describes a pharmaceutical compound (I) that
CC   contains an N-terminal glutamic acid (Glu) or glutamine (Gln) residue in
CC   the form of an addition salt with a strong, physiologically acceptable
CC   acid (II). Also described are: (a) a pharmaceutical composition
CC   containing at least one (I); (b) a vaccine containing at least one (I)
CC   where this is a major histocompatibility complex (MHC) ligand (Ia); (c) a
CC   method for in vitro diagnosis of diseases associated with the presence of
CC   (Ia); (d) a kit for method (c) that includes a (Ia); and (e) a process
CC   for preparing (I). (I) has immunomodulator, endocrine, antiallergic,
CC   neuroprotectant, virucidal, bactericidal, antiparasitic, fungicidal and
CC   cytostatic activities. (I) are useful, in human or veterinary medicine,
CC   in pharmaceutical compositions (for treating immune disorders, e.g.
CC   immune deficiency, autoimmune states, hypersensitivity, allergy, graft
CC   rejection, infection, hormonal disorders and central nervous system
CC   diseases), also, where (I) is a MHC ligand (Ia), in vaccines for
CC   treatment or prevention of: (i) viral, bacterial, parasitic or fungal
CC   infections; or (ii) of cancers. A particular application is in anti-
CC   melanoma vaccines. (I) are also useful for in vitro diagnosis of diseases
CC   associated with interactions between MHC and (I), e.g. melanoma and human
CC   immunodeficiency virus infection. AAM98898 to AAM99592 represent peptides
CC   which can be used in pharmaceutical compounds from the present invention
XX
SQ   Sequence 15 AA;
```

```
  Query Match             44.3%;  Score 35;  DB 4;  Length 15;
  Best Local Similarity   75.0%;  Pred. No. 23;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VHFFKDIS 12
              |||||:|:
Db          5 VHFFKNIA 12


RESULT 13
AAW72354
ID   AAW72354 standard; peptide; 7 AA.
XX
AC   AAW72354;
XX
DT   16-DEC-1998  (first entry)
XX
DE   Human myelin basic protein peptide formula.
XX
KW   Human; myelin basic protein; MBP; multiple sclerosis; anti-MBP; MS.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Val can be attached to: hydrogen; hydroxy; an
FT                   amino acid residue; and a polypeptide residue"
FT   Modified-site   7
FT                   /note= "Ile can be attached to: hydrogen; hydroxy; an
FT                   amino acid residue; and a polypeptide residue"
XX
PN   WO9845327-A1.
XX
PD   15-OCT-1998.
XX
PF   03-APR-1998;   98WO-CA000290.
XX
PR   04-APR-1997;   97CA-02201841.
XX
PA   (UYAL-) UNIV ALBERTA.
XX
PI   Warren KG,  Catz I;
XX
DR   WPI; 1998-568336/48.
XX
PT   Peptide and its derivatives for treatment of multiple sclerosis - is
PT   capable of neutralising or modulating production of anti-myelin basic
PT   protein.
XX
PS   Claim 4; Page 46; 75pp; English.
XX
CC   The present sequence represents a myelin basic protein (MBP) peptide
CC   formula. MBP peptides are capable of neutralising or modulating the
CC   production of anti-myelin basic protein. The present invention also
CC   describes a method for treating multiple sclerosis (MS). The method
CC   comprises administering to the patient an MBP peptide of the formula: R1-
CC   Val-His-Phe-Phe-Lys-Asn-Ile-R2 where R1, R2 = H, OH, or an amino acid
CC   residue and a polypeptide residue, provided that R1 and R2 are not both H
CC   or OH at the same time
XX
SQ   Sequence 7 AA;

  Query Match             43.0%;  Score 34;  DB 2;  Length 7;
  Best Local Similarity   85.7%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VHFFKDI 11
              |||||:|
Db          1 VHFFKNI 7


RESULT 14
AAR61755
ID   AAR61755 standard; peptide; 8 AA.
XX
AC   AAR61755;
XX
DT   25-MAR-2003  (revised)
DT   12-MAY-1995  (first entry)
XX
DE   MBP peptide 86, potential binder of HLA-A2.1.
XX
KW   antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV; HIV1;
KW   plasma specific antigen; hepatitis B virus; Epstein Barr;
KW   human immunodeficiency virus; human papilloma virus; p53; c-ERB2; MAGE-1;
KW   melanoma antigen-1; core antigen; surface antigen;
KW   pharmaceutical composition; in vivo; ex vivo; therapeutic; diagnostic;
KW   MHC class I molecule; major histocompatability complex; HLA-A2.1; 9mer;
KW   10mer; anchor; human leukocyte antigen; PLP; 8mer; algorithm prediction;
KW   MBP.
XX
OS   Homo sapiens.
XX
PN   WO9420127-A1.
XX
PD   15-SEP-1994.
XX
PF   04-MAR-1994;   94WO-US002353.
XX
PR   05-MAR-1993;   93US-00027146.
PR   04-JUN-1993;   93US-00073205.
PR   29-NOV-1993;   93US-00159184.
XX
PA   (CYTE-) CYTEL CORP.
XX
PI   Grey HM,  Sette A,  Sidney J,  Kast W;
XX
DR   WPI; 1994-302678/37.
XX
PT   Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for
PT   treatment or prophylaxis of cancer, virus infection or autoimmune
PT   diseases.
XX
PS   Disclosure; Page 117; 138pp; English.
XX
CC   AAR61714-837 are potential peptide binders of HLA-A2.1 motif. These
CC   peptides are thus potentially immunogenic. They were predicted by using
CC   an algorithm, which assigns a score for each amino acid, at each position
CC   along a peptide. A peptide is scored in the 'Grouped Ratio' algorithm as
CC   a product of the scores of each of its residues. This value can then be
CC   used to predict a population of peptides with the highest occurrence of
CC   good binders. The peptides of the invention can induce cytotoxic T
CC   lymphocytes which can react with target cells. They can be used for the
CC   treatment or prophlyaxis of cancer, eg. prostate cancer or lymphoma, etc.
CC   (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 8 AA;

  Query Match             43.0%;  Score 34;  DB 2;  Length 8;
  Best Local Similarity   85.7%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VHFFKDI 11
              |||||:|
Db          2 VHFFKNI 8


RESULT 15
ADR48840
ID   ADR48840 standard; peptide; 8 AA.
XX
AC   ADR48840;
XX
DT   04-NOV-2004  (first entry)
XX
```

DE Inhibitory compound #52.
XX
KW Major histocompatibility complex; MHC class II molecule;
KW T cell proliferation inhibition; leukocyte antigen-DR2;
KW autoimmune disease; multiple sclerosis; Behcet's disease;
KW Crohn's disease; glomerulonephritis; Grave's disease;
KW Guillain-Barre syndrome; Hashimoto's thyroiditis;
KW idiopathic thrombocytopenic purpura; lichen planus; lupus erythematosus;
KW Meniere's disease; type 1 diabetes; psoriasis; Reiter's syndrome;
KW rheumatoid arthritis; scleroderma; Sjogren's syndrome;
KW Wegener's granulomatosis; myasthenia gravis; ulcerative colitis; uveitis;
KW stiff-man syndrome; vasculitis.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "Acetylated residue"
XX
PN US2004162242-A1.
XX
PD 19-AUG-2004.
XX
PF 13-FEB-2004; 2004US-00778756.
XX
PR 14-FEB-2003; 2003US-0447949P.
XX
PA (OLSO/) OLSON G L.
PA (COOK/) COOK C M.
PA (SELF/) SELF C R.
XX
PI Olson GL, Cook CM, Self CR;
XX
DR WPI; 2004-634115/61.
XX
PT New amino acid derivatives, used as modulators of antigen presentation by
PT human leukocyte antigen-DR class II major histocompatibility complex
PT molecules, for treating autoimmune diseases.
XX
PS Claim 10; Page 33; 35pp; English.
XX
CC The invention relates to amino acid derivatives/compounds or their salts.
CC The invention also relates to a pharmaceutical composition comprising a
CC compound of the invention and a carrier. The compound or its salt is
CC useful for treating or preventing a disease responsive to the inhibition
CC of antigen binding to an MHC class II molecule, a disease responsive to
CC the inhibition of antigen presentation by an MHC class II molecule or a
CC disease responsive to the inhibition of T cell proliferation, which
CC involves administering a compound or its salt to a patient in need of
CC treatment, where the MHC class II molecule is human leukocyte antigen-
CC DR2. The compound or its salt is useful for treating or preventing
CC autoimmune diseases, such as multiple sclerosis, Behcet's disease,
CC Crohn's disease, glomerulonephritis, Grave's disease, Guillain-Barre
CC syndrome, Hashimoto's thyroiditis, idiopathic thrombocytopenic purpura,
CC lichen planus, lupus erythematosus, Meniere's disease, type 1 diabetes,
CC psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma,
CC Sjogren's syndrome, Wegener's granulomatosis, myasthenia gravis,
CC ulcerative colitis, uveitis, stiff-man syndrome and vasculitis. This
CC sequence represents an MHC class II antigen binding inhibitory compound
CC of the invention.
XX
SQ Sequence 8 AA;

Query Match 43.0%; Score 34; DB 8; Length 8;
Best Local Similarity 85.7%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VHFFKDI 11
||||:|
Db 2 VHFFKNI 8

Search completed: February 22, 2005, 09:24:50

Job time : 67.6667 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-12
Perfect score: 85
Sequence: 1 TGGVYHFVKKHVHES 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	28.2	13	2	PT0304	Ig heavy chain CRD
2	24	28.2	14	2	PC7079	unidentified 27.2K
3	23	27.1	14	2	S74128	superoxide dismuta
4	21	24.7	8	2	PL0184	capsid protein VP-
5	21	24.7	15	2	PN0662	dystrophin-associa
6	21	24.7	15	2	PQ0780	NADH2 dehydrogenas
7	20.5	24.1	14	2	PL0142	carbon-monoxide de
8	20	23.5	7	1	NYPG7	hypothalamic hepta
9	20	23.5	11	2	PQ0733	unidentified 6.0/1
10	20	23.5	14	2	S03530	Ig heavy chain J r
11	20	23.5	14	2	S57569	T cell receptor V-
12	20	23.5	14	2	S57638	T cell receptor V-
13	20	23.5	15	2	A61145	dihydrofolate redu
14	19.5	22.9	15	2	PA0093	enniatin synthetas
15	19	22.4	10	2	C61440	polygalacturonase
16	19	22.4	12	2	A33099	163K exoantigen -
17	19	22.4	12	2	PT0257	Ig heavy chain CRD
18	19	22.4	12	2	A34858	proteinase E - bla
19	19	22.4	14	2	PT0294	Ig heavy chain CRD
20	19	22.4	14	2	PH1601	Ig H chain V-D-J r
21	19	22.4	14	2	A47421	leukotriene B-4 12
22	19	22.4	15	2	S57201	basic proteinase I
23	19	22.4	15	2	PC1317	large granule L4 c
24	19	22.4	15	2	PH1613	Ig H chain V-D-J r
25	18	21.2	9	2	E28854	fibrinopeptide B -
26	18	21.2	10	2	S77980	cytochrome-c oxida
27	18	21.2	12	2	G64003	hypothetical prote
28	18	21.2	12	2	PH1464	T-cell receptor be
29	18	21.2	13	2	PT0263	Ig heavy chain CRD
30	18	21.2	13	2	S47380	T-cell antigen rec
31	18	21.2	13	2	AB0764	his operon leader
32	18	21.2	13	2	S60046	early nodulin 40 -
33	18	21.2	13	2	JQ2309	hypothetical 1.6K
34	18	21.2	13	2	JQ2319	hypothetical 1.6K
35	18	21.2	14	2	PT0254	Ig heavy chain CRD
36	18	21.2	14	2	PH0915	T-cell receptor be
37	18	21.2	14	2	S29632	xylan 1,4-beta-xyl
38	18	21.2	14	2	I54945	gene C protein - E
39	17	20.0	5	2	C53284	T-cell receptor be
40	17	20.0	6	2	JN0861	peptidyl-dipeptida
41	17	20.0	6	2	PT0723	T-cell receptor be
42	17	20.0	7	2	PT0719	T-cell receptor be
43	17	20.0	7	2	PT0702	T-cell receptor be
44	17	20.0	8	2	PT0595	T-cell receptor be
45	17	20.0	8	2	PT0530	T-cell receptor be

ALIGNMENTS

RESULT 1
PT0304
Ig heavy chain CRD3 region (clone 5-115B) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0304
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0304
A;Molecule type: DNA
A;Residues: 1-13 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 28.2%; Score 24; DB 2; Length 13;
Best Local Similarity 36.4%; Pred. No. 6.8e+02;
Matches 4; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy          3 GVYHFVKKHVH 13
              | | |  :  |
Db          3 GYYDFWSREAH 13
```

RESULT 2
PC7079
unidentified 27.2K protein - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C;Accession: PC7079
R;Tsugita, A.; Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Matsui, T.; Watanabe, Y.;
Electrophoresis 21, 1853-1871, 2000
A;Title: Proteome analysis of mouse brain: Two-dimensional electrophoresis profiles of t
A;Reference number: PC7072
A;Accession: PC7079
A;Molecule type: protein
A;Residues: 1-14 <TSU>
A;Cross-references: UNIPROT:Q8K3J1
A;Experimental source: strain C57BL/6Cr Slc, male; brain, striatum
C;Keywords: brain

Query Match 28.2%; Score 24; DB 2; Length 14;
Best Local Similarity 45.5%; Pred. No. 7.3e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

```
Qy          5 YHFVKKHVHES 15
              | :| |   ||
Db          2 YKYVNKKEQES 12
```

RESULT 3

S74128
superoxide dismutase (EC 1.15.1.1) 1 (Ni) - Streptomyces coelicolor (fragment)
C;Species: Streptomyces coelicolor
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999
C;Accession: S74128
R;Kim, E.J.; Kim, H.P.; Hah, Y.C.; Roe, J.H.
Eur. J. Biochem. 241, 178-185, 1996
A;Title: Differential expression of superoxide dismutases containing Ni and Fe/Zn in Str
A;Reference number: S74128; MUID:97054607; PMID:8898904
A;Accession: S74128
A;Molecule type: protein
A;Residues: 1-14 <KIM>
A;Experimental source: ATCC 10147
C;Function:
A;Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C;Keywords: metalloprotein; nickel; oxidoreductase; tetramer

Query Match 27.1%; Score 23; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GGVY 5
||||
Db 6 GGVY 9

RESULT 4
PL0184
capsid protein VP-1 - murine poliovirus (fragment)
C;Species: murine poliovirus, Theiler's encephalomyelitis virus
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 20-Feb-1995
C;Accession: PL0184
R;Zurbriggen, A.; Hogle, J.M.; Fujinami, R.S.
J. Exp. Med. 170, 2037-2049, 1989
A;Title: Alteration of amino acid 101 within capsid protein VP-1 changes the pathogenici
A;Reference number: PL0184; MUID:90063468; PMID:2479706
A;Accession: PL0184
A;Molecule type: genomic RNA
A;Residues: 1-8 <ZUR>
C;Keywords: capsid protein

Query Match 24.7%; Score 21; DB 2; Length 8;
Best Local Similarity 42.9%; Pred. No. 2.8e+05;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGGVYHF 7
:||: :|
Db 1 SGGITNF 7

RESULT 5
PN0662
dystrophin-associated glycoprotein A3a-I - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 07-May-1999
C;Accession: PN0662
R;Yoshida, M.; Mizuno, Y.; Nonaka, I.; Ozawa, E.
J. Biochem. 114, 634-639, 1993
A;Title: A dystrophin-associated glycoprotein, A3a (one of 43DAG doublets), is retained
A;Reference number: PN0662; MUID:94156881; PMID:8113213
A;Accession: PN0662
A;Molecule type: protein
A;Residues: 1-15 <YOS>
C;Comment: This protein is retained in Duchenne type muscular dystrophy muscle.
C;Keywords: glycoprotein; skeletal muscle

Query Match 24.7%; Score 21; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 2.5e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 FVKKHV 12
|:|| |
Db 5 FIKKGV 10

RESULT 6
PQ0780
NADH2 dehydrogenase (EC 1.6.99.3) 39K chain - fava bean mitochondrion (fragment)
N;Alternate names: complex I 39K chain; NADH-ubiquinone reductase 39K chain
C;Species: mitochondrion Vicia faba (fava bean)
C;Date: 03-May-1994 #sequence_revision 07-Oct-1994 #text_change 09-Jul-2004
C;Accession: PQ0780
R;Leterme, S.; Boutry, M.
Plant Physiol. 102, 435-443, 1993
A;Title: Purification and preliminary characterization of mitochondrial complex I (NADH:
A;Reference number: PQ0775; MUID:94151437; PMID:8108509
A;Accession: PQ0780
A;Molecule type: protein
A;Residues: 1-15 <LET>
A;Cross-references: UNIPROT:Q7M2F9
C;Comment: Complex I, mitochondrial NADH-ubiquiquinone reductase, is the first of the th
ranging from 5K to 75K.
C;Comment: This enzyme catalyzes electron transfer from endogenous NADH to ubiquinone by
C;Genetics:
A;Genome: mitochondrion
C;Keywords: electron transfer; mitochondrion; oxidoreductase

Query Match 24.7%; Score 21; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 2.5e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 GVYHFVK 9
|| | |:
Db 7 GVGHLVR 13

RESULT 7
PL0142
carbon-monoxide dehydrogenase (EC 1.2.99.2) medium chain - Pseudomonas carboxydoflava (f
C;Species: Pseudomonas carboxydoflava
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993
C;Accession: PL0142
R;Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.
Arch. Microbiol. 152, 335-341, 1989
A;Title: Homology and distribution of CO dehydrogenase structural genes in carboxydotrop
A;Reference number: PL0138; MUID:90055678; PMID:2818128
A;Accession: PL0142
A;Molecule type: protein
A;Residues: 1-14 <KRA>
C;Comment: Carbon-monoxide dehydrogenase consists of three polypeptide chains: large, me
C;Keywords: oxidoreductase

Query Match 24.1%; Score 20.5; DB 2; Length 14;
Best Local Similarity 62.5%; Pred. No. 2.8e+03;
Matches 5; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 5 YHFVKKHV 12
|| |||
Db 8 YH-APKHV 14

RESULT 8
NYPG7
hypothalamic heptapeptide - pig
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 01-Sep-1981 #sequence_revision 01-Sep-1981 #text_change 09-Jul-2004
C;Accession: A01417
R;Chang, R.C.C.; Huang, W.Y.; Arimura, A.; Redding, T.W.; Coy, D.H.; Saffran, M.; Kong,
Horm. Metab. Res. 13, 228-232, 1981
A;Title: Isolation, structure and synthesis of a heptapeptide with in vitro ACTH-releasi
A;Reference number: A01417; MUID:81213980; PMID:6263778
A;Accession: A01417
A;Molecule type: protein
A;Residues: 1-7 <CHA>
A;Cross-references: UNIPROT:P01153
C;Superfamily: hypothalamic heptapeptide

```
C;Keywords: hypothalamus

  Query Match             23.5%;  Score 20;  DB 1;  Length 7;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+05;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 VYHFVK 9
              :||  |
Db          2 IYHSYK 7


RESULT 9
PQ0733
unidentified 6.0/15K protein [imported] - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: PQ0733
R;Komatsu, S.; Kajiwara, H.; Hirano, H.
Theor. Appl. Genet. 86, 935-942, 1993
A;Title: A rice protein library; a data-file of rice proteins separated by two-dimension
A;Reference number: PQ0696
A;Accession: PQ0733
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <KOM>

  Query Match             23.5%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 HVH 13
              |||
Db          5 HVH 7


RESULT 10
S03530
Ig heavy chain J region (JH-4) - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 21-Jul-2000
C;Accession: S03530
R;Schwager, J.; Grossberger, D.; du Pasquier, L.
EMBO J. 7, 2409-2415, 1988
A;Title: Organization and rearrangement of immunoglobulin M genes in the amphibian Xenop
A;Reference number: S01158; MUID:89052653; PMID:2903824
A;Accession: S03530
A;Molecule type: DNA
A;Residues: 1-14 <SCH>
A;Cross-references: EMBL:X14918; NID:g64805; PIDN:CAA33043.1; PID:g1334657

  Query Match             23.5%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   75.0%;  Pred. No. 3.4e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 VYHF 7
              |||:
Db          1 VYHW 4


RESULT 11
S57569
T cell receptor V-J junctional alpha chain region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57569
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A;Reference number: S57494
A;Accession: S57569
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-14 <BUR>
A;Cross-references: EMBL:Z49955; NID:g887482; PIDN:CAA90226.1; PID:g887483
C;Keywords: T-cell receptor

  Query Match             23.5%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 3.4e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGGVYHF 7
              ||  ::|
Db          7 TGNQFYF 13


RESULT 12
S57638
T cell receptor V-J junctional alpha chain region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57638
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A;Reference number: S57494
A;Accession: S57638
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-14 <BUR>
A;Cross-references: EMBL:Z49964; NID:g886676; PIDN:CAA90238.1; PID:g886677
C;Keywords: T-cell receptor

  Query Match             23.5%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 3.4e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 TGGVYHF 7
              ||  ::|
Db          7 TGNQFYF 13


RESULT 13
A61145
dihydrofolate reductase (EC 1.5.1.3) - Mycobacterium smegmatis (fragment)
C;Species: Mycobacterium smegmatis
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 17-Mar-1999
C;Accession: A61145
R;Sirawaraporn, W.; Sirawaraporn, R.; Chanpongsri, A.; Jacobs Jr., W.R.; Santi, D.V.
Exp. Parasitol. 72, 184-190, 1991
A;Title: Purification and characterization of dihydrofolate reductase from wild-type and
A;Reference number: A61145; MUID:91184314; PMID:2009922
A;Accession: A61145
A;Molecule type: protein
A;Residues: 1-15 <SIR>
C;Keywords: NADP; oxidoreductase

  Query Match             23.5%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   75.0%;  Pred. No. 3.7e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TGGV 4
              |||:
Db         10 TGGI 13


RESULT 14
PA0093
enniatin synthetase - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0093
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotricl
A;Reference number: PA0051
A;Accession: PA0093
```

```
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Z6

  Query Match             22.9%;  Score 19.5;  DB 2;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 4.4e+03;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    3;  Gaps    1;

Qy          4 VYHFVKKHVH 13
              || ||   ||
Db          1 VYTFV---VH 7


RESULT 15
C61440
polygalacturonase (EC 3.2.1.15) II b - Aspergillus sp. (fragment)
C;Species: Aspergillus sp.
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 06-Dec-1996
C;Accession: C61440
R;Stratilova, E.; Markovic, O.; Skrovinova, D.; Rexova-Benkova, L.; Jornvall, H.
J. Protein Chem. 12, 15-22, 1993
A;Title: Pectinase Aspergillus sp. polygalacturonase: multiplicity, divergence, and stru
A;Reference number: A61440; MUID:93151962; PMID:8427629
A;Accession: C61440
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <STR>
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match             22.4%;  Score 19;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 3.4e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GGVYHF 7
              || : |
Db          2 GGAFTF 7


Search completed: February 22, 2005, 09:46:29
Job time : 12.1333 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-12
Perfect score: 85
Sequence: 1 TGGVYHFVKKHVHES 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	36.5	15	2	Q64FK1	Q64fk1 homo sapien
2	26	30.6	13	2	Q6LDM2	Q6ldm2 pseudomonas
3	26	30.6	13	2	Q7ZZN6	Q7zzn6 xenopus lae
4	25	29.4	11	2	Q9X9S6	Q9x9s6 streptomyce
5	24	28.2	13	1	GFA1_RAT	P82808 rattus norv
6	23	27.1	9	2	Q8QRR5	Q8qrr5 transmissib
7	23	27.1	9	2	Q8QRR6	Q8qrr6 transmissib
8	23	27.1	10	2	Q6R7V4	Q6r7v4 carlia zuma
9	23	27.1	11	2	Q68LE0	Q68le0 pyriglena l
10	23	27.1	11	2	Q68LE9	Q68le9 myrmotherul
11	23	27.1	12	1	UKA2_HUMAN	P31144 homo sapien
12	23	27.1	12	2	Q6AV52	Q6av52 oryza sativ
13	23	27.1	14	2	Q6R7V0	Q6r7v0 carlia viva
14	22	25.9	8	2	Q70Y68	Q70y68 prostanther
15	22	25.9	11	2	Q70Y70	Q70y70 thorncrofti
16	22	25.9	12	2	Q8GSB9	Q8gsb9 lolium pere
17	22	25.9	13	1	AU11_LITRA	P82386 litoria ran
18	22	25.9	13	1	AU12_LITRA	P82387 litoria ran
19	22	25.9	14	1	ADFA_TENMO	P82965 tenebrio mo
20	22	25.9	15	2	Q7RHY2	Q7rhy2 plasmodium
21	22	25.9	15	2	Q6ZZ74	Q6zz74 lychnis flo
22	22	25.9	15	2	Q9FS98	Q9fs98 silene laco
23	21.5	25.3	13	2	Q66NE7	Q66ne7 borrelia bu
24	21	24.7	8	1	ALL9_CARMA	P81812 carcinus ma
25	21	24.7	9	2	Q8W8X4	Q8w8x4 diadema mex
26	21	24.7	10	2	Q7RSJ0	Q7rsj0 plasmodium
27	21	24.7	13	2	Q6LCB1	Q6lcb1 rattus norv
28	21	24.7	13	2	Q90XG9	Q90xg9 gallus gall
29	21	24.7	14	2	Q86578	Q86578 sigma virus
30	21	24.7	15	1	UC14_MAIZE	P80620 zea mays (m
31	21	24.7	15	2	Q7M2F9	Q7m2f9 vicia faba
32	21	24.7	15	2	O68430	O68430 buchnera ap
33	20	23.5	7	1	HY7_PIG	P01153 sus scrofa
34	20	23.5	8	2	Q05403	Q05403 saccharomyc
35	20	23.5	8	2	Q6EGT1	Q6egt1 stemphylium
36	20	23.5	8	2	Q6EGV0	Q6egv0 stemphylium
37	20	23.5	8	2	Q6EGV3	Q6egv3 stemphylium
38	20	23.5	8	2	Q6EGW2	Q6egw2 stemphylium
39	20	23.5	12	2	Q9UC29	Q9uc29 homo sapien
40	20	23.5	12	2	Q8H6E6	Q8h6e6 hordeum vul
41	20	23.5	15	2	Q7R6W8	Q7r6w8 plasmodium
42	20	23.5	15	2	Q9TWF3	Q9twf3 dictyosteli
43	20	23.5	15	2	Q95762	Q95762 ctenosaura
44	20	23.5	15	2	O54298	O54298 salmonella
45	19.5	22.9	12	2	Q8CJ30	Q8cj30 mus musculu

ALIGNMENTS

RESULT 1
Q64FK1
ID Q64FK1 PRELIMINARY; PRT; 15 AA.
AC Q64FK1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Breast and ovarian cancer susceptibility protein 1 (Fragment).
GN Name=BRCA1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Soumittra N., Sridevi V., Shanta V., Rajkumar T.;
RT "Mutation analysis of BRCA1 gene.";
RL Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AY706913; AAU21565.1; -.
FT NON_TER 1 1
SQ SEQUENCE 15 AA; 1905 MW; 6E1A57496E31C01D CRC64;

Query Match 36.5%; Score 31; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy          4 VYHFVKKH 11
              || | :||
Db          3 VYKFARKH 10
```

RESULT 2
Q6LDM2
ID Q6LDM2 PRELIMINARY; PRT; 13 AA.
AC Q6LDM2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PAO;
RX MEDLINE=90036723; PubMed=2509433;
RA Hoshino T., Kose K.;
RT "Cloning and nucleotide sequence of braC, the structural gene for the
RT leucine-, isoleucine-, and valine-binding protein of Pseudomonas
RT aeruginosa PAO.";
RL J. Bacteriol. 171:6300-6306(1989).
DR EMBL; M31071; AAA88431.1; -.
KW Hypothetical protein.

```
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1648 MW;  F275BF6289EE3B51 CRC64;

  Query Match             30.6%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   30.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    6;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 VYHFVKKHVH 13
              :||:::: |:
Db          4 IYHYLQQLVN 13


RESULT 3
Q7ZZN6
ID   Q7ZZN6      PRELIMINARY;      PRT;    13 AA.
AC   Q7ZZN6;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Zinc finger protein (Fragment).
GN   Name=Zic3;
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Cerebellum;
RX   MEDLINE=22844363; PubMed=12963115; DOI=10.1016/S0925-4773(03)00082-0;
RA   Weber J.R., Sokol S.Y.;
RT   "Identification of a phylogenetically conserved activin-responsive
RT   enhancer in the Zic3 gene.";
RL   Mech. Dev. 120:955-964(2003).
DR   EMBL; AF506278; AAP20809.1; -.
FT   NON_TER       1      1
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1537 MW;  5DDA56257F6DF2C3 CRC64;

  Query Match             30.6%;  Score 26;  DB 2;  Length 13;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+03;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    2;  Gaps    1;

Qy          8 VKKH--VHES 15
              ::||  ||||
Db          1 LRKHMKVHES 10


RESULT 4
Q9X9S6
ID   Q9X9S6      PRELIMINARY;      PRT;    11 AA.
AC   Q9X9S6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein ORF9 (Fragment).
GN   Name=ORF9;
OS   Streptomyces lividans.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1916;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TK21;
RX   MEDLINE=99328982; PubMed=10400594;
RA   Martinez-Costa O.H., Martin-Triana A.J., Martinez E.,
RA   Fernandez-Moreno M.A., Malpartida F.;
RT   "An additinal regulatory gene for actinorhodin production in
RT   Streptomyces lividans involves a LysR-type transcriptional
RT   regulator.";
RL   J. Bacteriol. 181:4353-4364(1999).
DR   EMBL; Y18818; CAB51138.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   11 AA;  1160 MW;  D1BABA8EC1EDC412 CRC64;

  Query Match             29.4%;  Score 25;  DB 2;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          8 VKKHVH 13
              |: |||
Db          4 VRAHVH 9


RESULT 5
GFA1_RAT
ID   GFA1_RAT       STANDARD;      PRT;    13 AA.
AC   P82808;
DT   25-OCT-2004 (Rel. 45, Created)
DT   25-OCT-2004 (Rel. 45, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Glucosamine--fructose-6-phosphate aminotransferase [isomerizing] 1
DE   (EC 2.6.1.16) (Hexosephosphate aminotransferase 1) (D-fructose-6-
DE   phosphate amidotransferase 1) (GFAT 1) (Fragment).
GN   Name=Gfpt1;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE, FUNCTION, CATALYTIC ACTIVITY, AND ENZYME REGULATION.
RC   STRAIN=Sprague-Dawley; TISSUE=Liver;
RX   MEDLINE=20358980; PubMed=10898949; DOI=10.1006/abbi.2000.1895;
RA   Huynh Q.K., Gulve E.A., Dian T.;
RT   "Purification and characterization of glutamine:fructose 6-phosphate
RT   amidotransferase from rat liver.";
RL   Arch. Biochem. Biophys. 379:307-313(2000).
CC   -!- FUNCTION: Controls the flux of glucose into the hexosamine
CC       pathway. Most likely involved in regulating the availability of
CC       precursors for N- and O-linked glycosylation of proteins.
CC   -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-
CC       glutamate + D-glucosamine 6-phosphate.
CC   -!- ENZYME REGULATION: Inhibited by 4,4'-dithiodipyridine.
CC   -!- SUBUNIT: Homotetramer (Potential).
CC   -!- MISCELLANEOUS: Optimum pH is 7.5.
CC   -!- SIMILARITY: Contains 1 type-2 glutamine amidotransferase domain.
DR   MEROPS; C44.970; -.
DR   InterPro; IPR000583; GATase_2.
DR   PROSITE; PS00443; GATASE_TYPE_II; 1.
KW   Aminotransferase; Direct protein sequencing;
KW   Glutamine amidotransferase; Transferase.
FT   INIT_MET      0      0       By similarity.
FT   ACT_SITE      1      1       GATase (By similarity).
FT   DOMAIN        1    >13       Glutamine amidotransferase (Potential).
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1553 MW;  0944F6BB32DB473B CRC64;

  Query Match             28.2%;  Score 24;  DB 1;  Length 13;
  Best Local Similarity   30.0%;  Pred. No. 3.6e+03;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 GVYHFVKKHV 12
              |:: ::  ||
Db          2 GIFAYLNYHV 11


RESULT 6
Q8QRR5
ID   Q8QRR5      PRELIMINARY;      PRT;     9 AA.
AC   Q8QRR5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Spike glycoprotein (Fragment).
OS   Transmissible gastroenteritis virus.
```

```
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus; Group 1 species.
OX   NCBI_TaxID=11149;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KT2;
RA   Kim S.J., Kwon H.M.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF481368; AAL89748.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1087 MW;  34E3C1F2C33B1047 CRC64;

  Query Match             27.1%;  Score 23;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KKHVH 13
              | |||
Db          5 KVHVH 9


RESULT 7
Q8QRR6
ID   Q8QRR6      PRELIMINARY;      PRT;     9 AA.
AC   Q8QRR6;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Spike glycoprotein (Fragment).
OS   Transmissible gastroenteritis virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC   Coronaviridae; Coronavirus; Group 1 species.
OX   NCBI_TaxID=11149;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HKT2;
RA   Kim S.J., Kwon H.M.;
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF481367; AAL89746.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   9 AA;  1087 MW;  34E3C1F2C33B1047 CRC64;

  Query Match             27.1%;  Score 23;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KKHVH 13
              | |||
Db          5 KVHVH 9


RESULT 8
Q6R7V4
ID   Q6R7V4      PRELIMINARY;      PRT;    10 AA.
AC   Q6R7V4;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN   Name=GAPDH;
OS   Carlia zuma.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Scincoidea;
OC   Scincidae; Carlia.
OX   NCBI_TaxID=260893;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dolman G., Phillips B.;
RT   "Single copy nuclear DNA markers characterized for comparative
RT   phylogeography in Australian wet tropics rainforest skinks.";
RL   Mol. Ecol. Notes 4:185-187(2004).
DR   EMBL; AY508912; AAS09890.1; -.
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1171 MW;  9D0ABB2322C9C1EA CRC64;

  Query Match             27.1%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 HFVK 9
              ||||
Db          4 HFVK 7


RESULT 9
Q68LE0
ID   Q68LE0      PRELIMINARY;      PRT;    11 AA.
AC   Q68LE0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN   Name=NAPDH;
OS   Pyriglena leuconota.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Formicariidae;
OC   Pyriglena.
OX   NCBI_TaxID=183187;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15283860;
RA   Irestedt M., Fjeldsa J., Nylander J.A., Ericson P.G.;
RT   "Phylogenetic relationships of typical antbirds (Thamnophilidae) and
RT   test of incongruence based on Bayes factors.";
RL   BMC Evol. Biol. 4:23-23(2004).
DR   EMBL; AY677056; AAT96981.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1242 MW;  9D0AEB2622C9C1EA CRC64;

  Query Match             27.1%;  Score 23;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 4.5e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 HFVK 9
              ||||
Db          5 HFVK 8


RESULT 10
Q68LE9
ID   Q68LE9      PRELIMINARY;      PRT;    11 AA.
AC   Q68LE9;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN   Name=NAPDH;
OS   Myrmotherula fulviventris.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Passeriformes; Thamnophilidae;
OC   Myrmotherula.
OX   NCBI_TaxID=288045;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15283860;
RA   Irestedt M., Fjeldsa J., Nylander J.A., Ericson P.G.;
RT   "Phylogenetic relationships of typical antbirds (Thamnophilidae) and
RT   test of incongruence based on Bayes factors.";
RL   BMC Evol. Biol. 4:23-23(2004).
DR   EMBL; AY677047; AAT96972.1; -.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1242 MW;  9D0AEB2622C9C1EA CRC64;
```

```
  Query Match             27.1%;  Score 23;  DB 2;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 4.5e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 HFVK 9
              ||||
Db          5 HFVK 8


RESULT 11
UKA2_HUMAN
ID   UKA2_HUMAN     STANDARD;      PRT;    12 AA.
AC   P31144;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Unknown protein from 2D-PAGE of epidermal keratinocytes (Spot 1617)
DE   (Fragments).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Keratinocytes;
RX   MEDLINE=93162043; PubMed=1286667;
RA   Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA   Vandekerckhove J.;
RT   "Microsequences of 145 proteins recorded in the two-dimensional gel
RT   protein database of normal human epidermal keratinocytes.";
RL   Electrophoresis 13:960-969(1992).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 6.93, its MW is: 81.6 kDa.
DR   Aarhus/Ghent-2DPAGE; 1617; IEF.
KW   Direct protein sequencing.
FT   NON_TER       1      1
FT   NON_CONS      7      8
FT   UNSURE        8      8
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1351 MW;  D6CD4A5E75F2C1F6 CRC64;

  Query Match             27.1%;  Score 23;  DB 1;  Length 12;
  Best Local Similarity   42.9%;  Pred. No. 4.9e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 FVKKHVH 13
              :: :|||
Db          2 YLGQHVH 8


RESULT 12
Q6AV52
ID   Q6AV52      PRELIMINARY;      PRT;    12 AA.
AC   Q6AV52;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein OSJNBb0101N11.9.
GN   Name=OSJNBb0101N11.9;
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA   Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA   Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA   Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA   Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA   White O., Salzberg S.L., Fraser C.M.;
RT   "Oryza sativa chromosome 3 BAC OSJNBb0101N11 genomic sequence.";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Buell R.;
RL   Submitted (JUL-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC105747; AAT76999.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   12 AA;  1323 MW;  C9C78EF8E2573B5D CRC64;

  Query Match             27.1%;  Score 23;  DB 2;  Length 12;
  Best Local Similarity   57.1%;  Pred. No. 4.9e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 GGVYHFV 8
              || | :|
Db          3 GGAYLYV 9


RESULT 13
Q6R7V0
ID   Q6R7V0      PRELIMINARY;      PRT;    14 AA.
AC   Q6R7V0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Glyceraldehyde-3-phosphate dehydrogenase (Fragment).
GN   Name=GAPDH;
OS   Carlia vivax.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Scincoidea;
OC   Scincidae; Carlia.
OX   NCBI_TaxID=124122;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dolman G., Phillips B.;
RT   "Single copy nuclear DNA markers characterized for comparative
RT   phylogeography in Australian wet tropics rainforest skinks.";
RL   Mol. Ecol. Notes 4:185-187(2004).
DR   EMBL; AY508916; AAS09894.1; -.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1750 MW;  51923E7F6D0ABB23 CRC64;

  Query Match             27.1%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 5.8e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 HFVK 9
              ||||
Db          4 HFVK 7


RESULT 14
Q70Y68
ID   Q70Y68      PRELIMINARY;      PRT;     8 AA.
AC   Q70Y68;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Ribosomal protein (Fragment).
GN   Name=rps16;
OS   Prostanthera nivea (snowy mintbush).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Prostantheroideae; Westringieae;
OC   Prostanthera.
OX   NCBI_TaxID=38863;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15019625; DOI=10.1016/j.ympev.2003.08.002;
RA   Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Harley M.M.,
RA   Willis F., Simonds M.S.J., Powell M.P., Savolainen V.;
```

```
RT   "Phylogeny and evolution of basils and allies (Ocimeae, Labiatae)
RT   based on three plastid DNA regions.";
RL   Mol. Phylogenet. Evol. 31:277-299(2004).
DR   EMBL; AJ505403; CAD45523.1; -.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Ribosomal protein.
FT   NON_TER       1      1
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  838 MW;  C821F2C058786415 CRC64;

  Query Match             25.9%;  Score 22;  DB 2;  Length 8;
  Best Local Similarity   60.0%;  Pred. No. 1.6e+06;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GGVYH 6
              ||: |
Db          3 GGIVH 7


RESULT 15
Q70Y70
ID   Q70Y70      PRELIMINARY;      PRT;    11 AA.
AC   Q70Y70;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Ribosomal protein (Fragment).
GN   Name=rps16;
OS   Thorncroftia longiflora.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Lamiales; Lamiaceae; Nepetoideae; Ocimeae; Thorncroftia.
OX   NCBI_TaxID=204202;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   PubMed=15019625; DOI=10.1016/j.ympev.2003.08.002;
RA   Paton A., Springate D.A., Sudde S., Otieno D., Grayer R., Harley M.M.,
RA   Willis F., Simonds M.S.J., Powell M.P., Savolainen V.;
RT   "Phylogeny and evolution of basils and allies (Ocimeae, Labiatae)
RT   based on three plastid DNA regions.";
RL   Mol. Phylogenet. Evol. 31:277-299(2004).
DR   EMBL; AJ505401; CAD45521.1; -.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
KW   Ribosomal protein.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1201 MW;  396BB78821F2C058 CRC64;

  Query Match             25.9%;  Score 22;  DB 2;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 6.7e+03;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GGVYH 6
              ||: |
Db          3 GGIVH 7


Search completed: February 22, 2005, 09:38:02
Job time : 53.6667 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-12
Perfect score: 85
Sequence: 1 TGGVYHFVKKHVHES 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	85	100.0	15	2	AAW04852	Aaw04852 Internal
2	37	43.5	13	5	ABB74425	Abb74425 Karyophil
3	36	42.4	15	5	ABJ10692	Abj10692 Japanese
4	36	42.4	15	5	ABJ10696	Abj10696 Japanese
5	35	41.2	15	2	AAW04855	Aaw04855 Internal
6	32	37.6	12	2	AAW80378	Aaw80378 Peptide e
7	32	37.6	15	5	AAU78249	Aau78249 Polypepti
8	31	36.5	7	4	AAU03943	Aau03943 Human GDN
9	31	36.5	13	7	ADB81184	Adb81184 Human pan
10	30	35.3	11	2	AAW11603	Aaw11603 Linear Zn
11	30	35.3	12	8	ADP83191	Adp83191 ICAM-5 FG
12	30	35.3	13	8	ADI32280	Adi32280 Cosmetic
13	30	35.3	15	3	AAB13485	Aab13485 C. tracho
14	30	35.3	15	4	AAG83161	Aag83161 Chlamydia
15	30	35.3	15	5	ABB94132	Abb94132 Chlamydia
16	30	35.3	15	6	ABR32837	Abr32837 Human can
17	30	35.3	15	6	ABR32981	Abr32981 Human can
18	30	35.3	15	6	ABR33030	Abr33030 Human can
19	30	35.3	15	6	ABR32938	Abr32938 Human can
20	30	35.3	15	6	ABR33038	Abr33038 Human can
21	29	34.1	10	6	ABJ20765	Abj20765 162P1E6 c
22	29	34.1	10	6	ABJ23568	Abj23568 162P1E6 c
23	29	34.1	10	6	ABJ22816	Abj22816 162P1E6 c
24	29	34.1	10	6	ABJ24215	Abj24215 162P1E6 c
25	29	34.1	10	6	ABJ21424	Abj21424 162P1E6 c
26	29	34.1	10	6	ABJ24945	Abj24945 162P1E6 c
27	29	34.1	10	6	ABJ22137	Abj22137 162P1E6 c
28	29	34.1	12	2	AAR63394	Aar63394 Peptide f
29	29	34.1	12	2	AAR96448	Aar96448 Hepatitis
30	29	34.1	12	2	AAR96449	Aar96449 Hepatitis
31	29	34.1	14	3	AAY78374	Aay78374 H. pylori
32	29	34.1	15	2	AAY30497	Aay30497 Mutant se
33	29	34.1	15	5	AAU76497	Aau76497 Ribosomal
34	29	34.1	15	8	ADP26532	Adp26532 Plasmodiu
35	28	32.9	9	2	AAY45437	Aay45437 Immunogen
36	28	32.9	9	4	AAB98468	Aab98468 HLA class
37	28	32.9	9	4	AAB96047	Aab96047 HLA-A11 b
38	28	32.9	9	4	AAG84524	Aag84524 Human leu
39	28	32.9	9	4	AAU06283	Aau06283 Human Leu
40	28	32.9	9	4	AAG88276	Aag88276 Human leu
41	28	32.9	9	4	ABP11505	Abp11505 HLA class
42	28	32.9	9	4	AAG89372	Aag89372 Human leu
43	28	32.9	9	4	AAJ00060	Aaj00060 Hepatitis
44	28	32.9	9	5	ABJ05792	Abj05792 Hepatitis
45	28	32.9	9	8	ADK39622	Adk39622 Hepatitis

ALIGNMENTS

RESULT 1
AAW04852
ID AAW04852 standard; peptide; 15 AA.
XX
AC AAW04852;
XX
DT 16-OCT-2003 (revised)
DT 18-FEB-1997 (first entry)
XX
DE Internal fragment of Epstein-Barr virus DNA polymerase.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Human herpesvirus 4.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 2; Page 45; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with

CC the autoimmune disease. This peptide is an internal peptide of Epstein-
CC Barr virus DNA polymerase and is implicated as a foreign epitope involved
CC in the aetiology or in remissions of multiple sclerosis. It has been
CC shown capable of inducing the proliferation of autoreactive T-cell clones
CC isolated from multiple sclerosis patients. (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 85; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGVYHFVKKHVHES 15
||||||||||||||
Db 1 TGGVYHFVKKHVHES 15

RESULT 2
ABB74425
ID ABB74425 standard; peptide; 13 AA.
XX
AC ABB74425;
XX
DT 18-APR-2002 (first entry)
XX
DE Karyophilic peptide SEQ ID NO:189.
XX
KW Fusogenic; nuclear localisation signal; NLS; encapsulation; lipogene;
KW liposome; micelle; karyophilic; cytostatic; antitumour; solid tumour;
KW peptide-lipid-polynucleotide complex; neoplastic disease; gene therapy;
KW breast carcinoma; prostate carcinoma.
XX
OS Schizosaccharomyces pombe.
XX
PN WO200193836-A2.
XX
PD 13-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US018657.
XX
PR 09-JUN-2000; 2000US-0210925P.
XX
PA (BOUL/) BOULIKAS T.
XX
PI Boulikas T;
XX
DR WPI; 2002-164295/21.
XX
PT Encapsulation of plasmid DNA (Lipogenes) and therapeutic agents with
PT nuclear localization signal/fusogenic peptide conjugates into targeted
PT liposome complexes.
XX
PS Claim 14; Page 66; 107pp; English.
XX
CC The present invention describes a method for producing micelles with
CC entrapped therapeutic agents. The method comprises: (1) combining
CC negatively charged agent with a cationic lipid in a ratio where 30-90 %
CC of the negatively charged atoms are neutralised by positive charges on
CC lipid molecules to form an electrostatic micelle complex in 20-80 %
CC ethanol; and (2) combining the micelle complex of (a) with fusogenic-
CC karyophilic peptide conjugates in a 0.0-0.3 ratio, therefore producing
CC micelles with entrapped therapeutic agents. Also described is a method
CC for delivering a therapeutic agent in vivo, comprising the administration
CC of the micelle. ABB74256 to ABB74858 represent specifically claimed
CC nuclear localisation signal (NLS) peptides for use in the method as the
CC fusogenic-karyophilic peptides. The micelles produced can have cytostatic
CC and antitumour activities. The peptide-lipid-polynucleotide complexes
CC produced are useful for inhibiting the progression of neoplastic
CC diseases. The invention relates to the field of gene therapy and is
CC directed toward methods for producing peptide-lipid-polynucleotide
CC complexes suitable for delivery of polynucleotides. The encapsulated
CC molecules display therapeutic efficacy in eradicating solid tumours
CC including but not limited to breast carcinoma or prostate carcinoma.
CC ABB74235 to ABB74255 are used in the exemplification of the present
CC invention
XX
SQ Sequence 13 AA;

Query Match 43.5%; Score 37; DB 5; Length 13;
Best Local Similarity 40.0%; Pred. No. 24;
Matches 4; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGVYHFVKKH 11
| ::|:::||
Db 2 GELFHYIRKH 11

RESULT 3
ABJ10692
ID ABJ10692 standard; peptide; 15 AA.
XX
AC ABJ10692;
XX
DT 28-NOV-2002 (first entry)
XX
DE Japanese encephalitis virus vaccine related peptide #87.
XX
KW Virucide; Japanese encephalitis viral infection; protective immunity;
KW chimeric synthetic peptide; viral infection; envelope glycoprotein;
KW vaccine; neutralising antibody.
XX
OS Unidentified.
OS Chimeric.
XX
PN WO200253182-A1.
XX
PD 11-JUL-2002.
XX
PF 04-JAN-2002; 2002WO-IN000003.
XX
PR 05-JAN-2001; 2001IN-DE000013.
XX
PA (BIOT-) SEC DEPT BIOTECHNOLOGY.
PA (NAVI-) NAT INST VIROLOGY.
PA (UYPU-) UNIV PUNE.
XX
PI Gore MM, Kolaskar AS, Dewasthaly SS, Kale UDK;
XX
DR WPI; 2002-548046/58.
XX
PT Vaccine composition for inducing protective immunity against Japanese
PT encephalitis viral infection, has synthetic peptide e.g. envelope
PT glycoprotein of virus, or neutralizing antibody inducing glycoprotein
PT sequences.
XX
PS Disclosure; Page 28; 44pp; English.
XX
CC The invention relates to a vaccine composition for humans and animals
CC against Japanese encephalitis viral infection. Tbe composition comprises
CC a chimeric synthetic peptide, especially an envelope glycoprotein of the
CC Japanese encephalitis virus, to induce protective immunity against the
CC viral infection, or a neutralising antibody inducing peptide sequences
CC from the envelope glycoproteins of Japanese encephalitis virus. The
CC vaccine composition is useful for inducing a protective immunity against
CC Japanese encephalitis viral infection. This sequence represents a
CC Japanese encephalitis virus vaccine peptide of the invention
XX
SQ Sequence 15 AA;

Query Match 42.4%; Score 36; DB 5; Length 15;
Best Local Similarity 46.2%; Pred. No. 41;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 2 GGVYHFVKKHVHE 14
|||:: : | ||:

```
Db          3 GGVFNSIGKAVHQ 15


RESULT 4
ABJ10696
ID   ABJ10696 standard; peptide; 15 AA.
XX
AC   ABJ10696;
XX
DT   28-NOV-2002  (first entry)
XX
DE   Japanese encephalitis virus vaccine related peptide #91.
XX
KW   Virucide; Japanese encephalitis viral infection; protective immunity;
KW   chimeric synthetic peptide; viral infection; envelope glycoprotein;
KW   vaccine; neutralising antibody.
XX
OS   Unidentified.
OS   Chimeric.
XX
PN   WO200253182-A1.
XX
PD   11-JUL-2002.
XX
PF   04-JAN-2002; 2002WO-IN000003.
XX
PR   05-JAN-2001; 2001IN-DE000013.
XX
PA   (BIOT-) SEC DEPT BIOTECHNOLOGY.
PA   (NAVI-) NAT INST VIROLOGY.
PA   (UYPU-) UNIV PUNE.
XX
PI   Gore MM,  Kolaskar AS,  Dewasthaly SS,  Kale UDK;
XX
DR   WPI; 2002-548046/58.
XX
PT   Vaccine composition for inducing protective immunity against Japanese
PT   encephalitis viral infection, has synthetic peptide e.g. envelope
PT   glycoprotein of virus, or neutralizing antibody inducing glycoprotein
PT   sequences.
XX
PS   Disclosure; Page 29; 44pp; English.
XX
CC   The invention relates to a vaccine composition for humans and animals
CC   against Japanese encephalitis viral infection. Tbe composition comprises
CC   a chimeric synthetic peptide, especially an envelope glycoprotein of the
CC   Japanese encephalitis virus, to induce protective immunity against the
CC   viral infection, or a neutralising antibody inducing peptide sequences
CC   from the envelope glycoproteins of Japanese encephalitis virus. The
CC   vaccine composition is useful for inducing a protective immunity against
CC   Japanese encephalitis viral infection. This sequence represents a
CC   Japanese encephalitis virus vaccine peptide of the invention
XX
SQ   Sequence 15 AA;

  Query Match             42.4%;  Score 36;  DB 5;  Length 15;
  Best Local Similarity   46.2%;  Pred. No. 41;
  Matches    6;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 GGVYHFVKKHVHE 14
              |||:: : | ||:
Db          3 GGVFNSIGKAVHQ 15


RESULT 5
AAW04855
ID   AAW04855 standard; peptide; 15 AA.
XX
AC   AAW04855;
XX
DT   18-FEB-1997  (first entry)
XX
DE   Internal fragment of herpes simplex virus DNA polymerase.
XX
KW   Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW   HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW   desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW   phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW   DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS   Herpes simplex virus.
XX
PN   WO9627387-A1.
XX
PD   12-SEP-1996.
XX
PF   07-MAR-1996;   96WO-US003182.
XX
PR   07-MAR-1995;   95US-00400796.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Strominger JL,  Wucherpfennig KW;
XX
DR   WPI; 1996-425218/42.
XX
PT   Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT   - useful in disease treatment, and method for identification of other
PT   self and non-self antigens implicated in auto-immune disease.
XX
PS   Claim 2; Page 47; 58pp; English.
XX
CC   Pharmaceutical preparations for tolerisation to antigens comprise either
CC   an isolated human non-collagen or non-myslin basic protein (MBP)
CC   polypeptide which is capable of tolerising an individual to an
CC   autoantigen; or an isolated human pathogen polypeptide capable of
CC   tolerising an individual to that polypeptide. In both cases, the
CC   polypeptide (whether self or non-self) includes an amino acid sequence
CC   corresponding to a sequence motif for a MHC class II protein, such as HLA
CC   -DR, which is associated with a human autoimmune disease and which binds
CC   to the polypeptide to activate autoreactive T-cells in individuals with
CC   the autoimmune disease. This peptide is an internal peptide of herpes
CC   simplex virus DNA polymerase and is implicated as a foreign epitope
CC   involved in the aetiology or in remissions of multiple sclerosis. It has
CC   been shown capable of inducing the proliferation of autoreactive T-cell
CC   clones isolated from multiple sclerosis patients
XX
SQ   Sequence 15 AA;

  Query Match             41.2%;  Score 35;  DB 2;  Length 15;
  Best Local Similarity   77.8%;  Pred. No. 60;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 FVKKHVHES 15
              ||| || ||
Db          7 FVKAHVRES 15


RESULT 6
AAW80378
ID   AAW80378 standard; peptide; 12 AA.
XX
AC   AAW80378;
XX
DT   14-JAN-1999  (first entry)
XX
DE   Peptide eluted after biopanning against maltose binding protein.
XX
KW   Intervening protein sequence; IVPS; protein splicing; protein production;
KW   maltose binding protein.
XX
OS   Synthetic.
XX
PN   US5834247-A.
XX
PD   10-NOV-1998.
```

```
XX
PF   05-MAR-1997;   97US-00811492.
XX
PR   09-DEC-1992;   92US-00004139.
PR   03-NOV-1993;   93US-00146885.
PR   28-JUN-1995;   95US-00496247.
PR   29-DEC-1995;   95US-00580555.
XX
PA   (NEWE ) NEW ENGLAND BIOLABS INC.
XX
PI   Hodges RA,  Perler FB,  Comb DG,  Southworth M,  Adam E,  Noren CJ;
PI   Xu M,  Chong SSC,  Jack WE;
XX
DR   WPI; 1999-008713/01.
XX
PT   New modified target proteins - which have controllable intervening
PT   protein sequence which can facilitate production, purification, labelling
PT   or isolation of target proteins.
XX
PS   Example 22; Fig 36; 123pp; English.
XX
CC   AAW80372-93 represent peptides eluted after biopanning against maltose
CC   binding protein, in the course of the invention. The specification
CC   describes IVPS (intervening protein sequence) regions which encode
CC   peptides which are removed via protein splicing to form the native
CC   protein. The specification describes a modified protein comprising a
CC   target protein or portion, fused either internally or terminally, to a
CC   IVPS, or to an amino- or carboxyl-terminal element of a IVPS. The IVPS
CC   are capable of excision from or cleavage of the modified protein upon
CC   predetermined conditions, in cis or trans, e.g. temparature increase,
CC   deglycosylation, unblocking of amino acid residues, treatment with
CC   chemical reagents. The methods can be used for modifying, producing,
CC   purifying, labelling or isolating target proteins such as enzymes,
CC   toxins, cytokines, glycoproteins and growth factors
XX
SQ   Sequence 12 AA;

  Query Match             37.6%;  Score 32;  DB 2;  Length 12;
  Best Local Similarity   62.5%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 HFVKKHVH 13
              | |: |||
Db          5 HHVRPHVH 12


RESULT 7
AAU78249
ID   AAU78249 standard; peptide; 15 AA.
XX
AC   AAU78249;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Polypeptide-calcium ion binding 14 N-terminus.
XX
KW   Polypeptide-calcium ion binding 14; cancer; HIV;
KW   human immunodeficiency virus infection.
XX
OS   Unidentified.
XX
PN   CN1327999-A.
XX
PD   26-DEC-2001.
XX
PF   12-JUN-2000; 2000CN-00116458.
XX
PR   12-JUN-2000; 2000CN-00116458.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-270051/32.
XX
PT   New polypeptide-calcium ion binding 14 and polynucleotide encoding it,
PT   for treating diseases such as cancer, and human immunodeficiency virus
PT   infection.
XX
PS   Example 5; Page 19 (Disclosure); 33pp; Chinese.
XX
CC   The invention relates to a new polypeptide-calcium ion binding 14, the
CC   polynucleotide encoding it, preparing the polypeptide by DNA
CC   recombination, application of the polypeptide in treating diseases such
CC   as cancer, human immunodeficiency virus (HIV) infection, the antagonist
CC   to the polynucleotide and its medical action, and the application of the
CC   polynucleotide are disclosed. The present sequence represents the
CC   polypeptide-calcium ion binding 14 N-terminus, used in an ELISA (enzyme-
CC   linked immunosorbent assay) experiment
XX
SQ   Sequence 15 AA;

  Query Match             37.6%;  Score 32;  DB 5;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 1.9e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 HFVKKHVH 13
              ||  ||:|
Db          3 HFNIKHIH 10


RESULT 8
AAU03943
ID   AAU03943 standard; peptide; 7 AA.
XX
AC   AAU03943;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Human GDNF F2c region peptide.
XX
KW   Persephin; F2a; F2c; GDNF; neurturin; artemin; human; mouse; rat; AIDS;
KW   growth factor receptor alpha1-RET protein tyrosine kinase; GFRalpha1-RET;
KW   trophic support; peripheral neuropathy; amyotrophic lateral sclerosis;
KW   Alzheimer's disease; Parkinson's disease; Huntington's disease; diabetes;
KW   acquired immunodeficiency syndrome; ischaemic stroke; acute brain injury;
KW   acute spinal cord injury; multiple sclerosis; nervous system tumour;
KW   neuroblastoma; enteric disease; idiopathic constipation; eosinopaenia;
KW   basopaenia; lymphopaenia; monocytopaenia; neutropaenia; anaemia;
KW   cardiac muscle degeneration; congestive heart failure; thrombocytopaenia.
XX
OS   Homo sapiens.
XX
PN   WO200147946-A2.
XX
PD   05-JUL-2001.
XX
PF   21-DEC-2000; 2000WO-US034852.
XX
PR   28-DEC-1999;   99US-00473551.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Milbrandt JD,  Baloh RH;
XX
DR   WPI; 2001-425618/45.
XX
PT   New growth factor that activates growth factor receptor alpha1-Ret
PT   protein-tyrosine kinase, for providing trophic support to a mammalian
PT   cell and producing differentiation of a mammalian cell in a patient.
XX
PS   Claim 6; Page 47; 73pp; English.
XX
CC   The sequence represents a human GDNF F2c region peptide. Residues from
CC   the F2a and/or F2c regions can substitute those of the F2a and F2c
CC   regions of the growth factor protein persephin. The substitutions can be
```

PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 131; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 9 AA;

Query Match 35.1%; Score 27; DB 4; Length 9;
Best Local Similarity 71.4%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 GGRRLFF 7
              ||| | |
Db          3 GGRHLIF 9
```

Search completed: February 22, 2005, 09:24:57
Job time : 67.6667 secs

```
CC   neutralising monoclonal antibody (588 antibody) that binds gp120 and also
CC   inhibits the binding of HIV to CD4. They were originally identified in
CC   the screening of phage epitope libraries having randomly generated
CC   epitope polypeptides accessible to the antibody. The libraries were ALPHA
CC   and EPSILON described in AAR83295 and AAR83299 respectively.The sequences
CC   of these polypeptides were deduced from their corresponding DNA sequence,
CC   determined by PCR
XX
SQ   Sequence 15 AA;

  Query Match             36.4%;  Score 28;  DB 2;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 5.9e+02;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 FVKAHVRE 14
              ||:|| |:
Db          8 FVRAHGRD 15


RESULT 13
ABG76687
ID   ABG76687 standard; peptide; 15 AA.
XX
AC   ABG76687;
XX
DT   05-NOV-2002  (first entry)
XX
DE   Hepatitis C virus (HCV)-specific ligand #18.
XX
KW   Binding specificity; anti-antigen antibody; serum; ADAM;
KW   antigen detection by antigen mimic; Hepatitis C virus infection; HCV;
KW   HCV-specific ligand; human.
XX
OS   Homo sapiens.
XX
PN   WO200237115-A1.
XX
PD   10-MAY-2002.
XX
PF   03-NOV-2000; 2000WO-IT000442.
XX
PR   03-NOV-2000; 2000WO-IT000442.
XX
PA   (KENT-) KENTON SRL.
XX
PI   Felici F,  Gargano N,  Minenkova O,  Monaci P;
XX
DR   WPI; 2002-599299/64.
XX
PT   Making diagnosis of antigen, by identifying binding specificity of anti-
PT   antigen antibody molecules in serum by antibody detection by antigen
PT   mimics methodology, and identifying antibodies associated with antigen.
XX
PS   Claim 26; Page 42; 86pp; English.
XX
CC   The present invention relates to a method for making a diagnosis of an
CC   antigen. The method involves identifying the binding specificity of the
CC   anti-antigen antibody molecules in serum by the antibody detection by
CC   antigen mimics (ADAM) methodology. The method comprises screening phage
CC   libraries using sera from antigen-infected and non-infected individuals,
CC   and identifying peptides binding antibodies (ligands) specifically
CC   associated with the antigen. The method of the invention can be used for
CC   the detection of infectious agents, particularly Hepatitis C virus (HCV).
CC   The invention provides HCV-specific ligands which are useful for the
CC   preparation of a diagnostic assay for detecting HCV infection in a
CC   subject. The HCV-specific ligands are also useful for the preparation of
CC   vaccines against HCV. ABG76670-ABG76743 represent HCV-specific ligands
CC   identified from human sera
XX
SQ   Sequence 15 AA;

  Query Match             36.4%;  Score 28;  DB 5;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 5.9e+02;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 RLFFVKAH 11
              ||: :|||
Db          3 RLWLMKAH 10


RESULT 14
AAJ02281
ID   AAJ02281 standard; peptide; 8 AA.
XX
AC   AAJ02281;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Hepatitis C virus epitope #2272.
XX
KW   Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW   antiviral.
XX
OS   Hepatitis C virus.
XX
PN   WO200121189-A1.
XX
PD   29-MAR-2001.
XX
PF   19-JUL-2000; 2000WO-US019774.
XX
PR   19-JUL-1999;   99US-00357737.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S,  Livingston BD,  Chesnut R;
PI   Baker DM,  Celis E,  Kubo RT,  Grey HM;
XX
DR   WPI; 2001-308046/32.
XX
PT   A new composition useful as a vaccines against hepatitis C virus.
XX
PS   Disclosure; Page 157; 214pp; English.
XX
CC   The present invention describes a composition comprising a prepared
CC   hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC   These are derived from HCV HLA-binding motifs. They are useful in
CC   vaccines for the prevention and treatment of HCV infection in humans. The
CC   present sequence is an epitope used in the disclosure of the invention
XX
SQ   Sequence 8 AA;

  Query Match             35.1%;  Score 27;  DB 4;  Length 8;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGRRLFF 7
              ||| | |
Db          2 GGRHLIF 8


RESULT 15
AAJ01219
ID   AAJ01219 standard; peptide; 9 AA.
XX
AC   AAJ01219;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Hepatitis C virus epitope #1210.
XX
KW   Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW   antiviral.
XX
OS   Hepatitis C virus.
XX
```

PS Claim 2; Page 2995; 3148pp; English.
XX
CC This invention describes novel antibodies that immunospecifically bind to
CC B Lymphocyte Stimulator (BLyS) polypeptides. BLyS is a member of the
CC tumour necrosis factor (TNF) super family and induces B cell
CC proliferation and differentiation. The antibodies of the invention have
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,
CC antirheumatic and antiAIDS activity and can be used in vaccines to
CC inhibit the expression and activity of BLyS. The antibodies bind to BLyS
CC and so may be used to detect and quantitate the presence of BLyS in
CC biological samples and may be used in this way to diagnose disease
CC associated with aberrant expression of BLyS. They may also be
CC administered to treat diseases associated with aberrant BLyS expression
CC and activity such as cancer, immune, and autoimmune disorders and
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
CC the antibodies and fragments of the antibodies described in the method of
CC the invention
XX
SQ Sequence 14 AA;

Query Match 36.4%; Score 28; DB 5; Length 14;
Best Local Similarity 55.6%; Pred. No. 5.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LFFVKAHVR 13
| | :||:|
Db 5 LLFPRAHLR 13

RESULT 11
ADG97285
ID ADG97285 standard; peptide; 14 AA.
XX
AC ADG97285;
XX
DT 11-MAR-2004 (first entry)
XX
DE scFV VHCDR3 peptide that immunospecifically binds BLyS SeqID 2469.
XX
KW antibody; B lymphocyte stimulator; BLyS; tumour necrosis factor;
KW B cell proliferation; differentiation; scFv; myasthenia gravis;
KW multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;
KW carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;
KW antiinflammatory; antiasthmatic; antiallergic; cytostatic.
XX
OS Unidentified.
XX
PN WO2003055979-A2.
XX
PD 10-JUL-2003.
XX
PF 14-NOV-2002; 2002WO-US036496.
XX
PR 16-NOV-2001; 2001US-0331469P.
PR 19-DEC-2001; 2001US-0340817P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;
XX
DR WPI; 2003-505530/47.
XX
PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator
PT (BLys), useful for detecting and treating diseases or disorders e.g.
PT rheumatoid arthritis, asthma and leukemia.
XX
PS Example 1; SEQ ID NO 2469; 394pp; English.
XX
CC This invention relates to novel antibodies that immunospecifically bind
CC to B lymphocyte stimulator (BLyS). The BLyS gene has been mapped to
CC chromosome 13q34 and encodes a protein that is a member of the tumour
CC necrosis factor superfamily and induces both in vivo and in vitro B cell
CC proliferation and differentiation. Specifically, it refers to single
CC chain antibody molecules (scFvs) derived, preferably, from the variable
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or
CC fragment thereof, of either human, murine, rat or monkey BLyS. The
CC present invention refers to the use of such antibodies in various methods
CC for the detection, diagnosis and prognosis of diseases related to the
CC aberrant expression or inappropriate function of BLyS or its receptor. As
CC such, these compositions are useful for identifying immune disorders
CC including myasthenia gravis and multiple sclerosis, inflammatory
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such
CC as AIDS and proliferative disorders including leukaemia, carcinoma and
CC lymphoma. Accordingly, they can be described as exhibiting various
CC activities such as antirheumatic, antiarthritic, neuroprotective,
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This
CC peptide sequence is a single chain antibody variable heavy CDR3 peptide
CC that immunospecifically binds BLyS of the invention.
XX
SQ Sequence 14 AA;

Query Match 36.4%; Score 28; DB 7; Length 14;
Best Local Similarity 55.6%; Pred. No. 5.5e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 5 LFFVKAHVR 13
| | :||:|
Db 5 LLFPRAHLR 13

RESULT 12
AAR93674
ID AAR93674 standard; peptide; 15 AA.
XX
AC AAR93674;
XX
DT 27-APR-1996 (first entry)
XX
DE HIV principal neutralisation epitope binding to 588 antibody.
XX
KW SPNE; selected principal neutralisation epitope; vaccine; HIV;
KW outer membrane proteosome; Neisseria; OMPC; AIDS; 588 antibody.
XX
OS Synthetic.
XX
PN GB2282381-A.
XX
PD 05-APR-1995.
XX
PF 23-SEP-1994; 94GB-00019257.
XX
PR 30-SEP-1993; 93US-00129723.
XX
PA (MERI) MERCK & CO INC.
XX
PI Arnold BA, Conley AJ, Keller PM;
XX
DR WPI; 1995-125268/17.
XX
PT New antigenic conjugate useful as vaccine for AIDS - comprising HIV
PT principal neutralisation epitope covalently linked to outer membrane
PT proteosome of Neisseria.
XX
PS Claim 14; Page 9; 69pp; English.
XX
CC An antigenic conjugate, useful as a vaccine for AIDS, has the formula
CC (SPNE)n-(OMPC), where SPNE is a selected principal neutralisation epitope
CC of HIV, which is one of 17 specified polypeptides (including the present
CC sequence) or their fragments containing at least 5 amino acids and
CC including the EWG region or its homologue in the sequence; OMPC is
CC purified outer membrane proteosome of Neisseria (pref. N. meningitidis);
CC and n is 1-200, indicating the number of SPNE moieties covalently linked
CC to the OMPC. The conjugates may be substituted by anions, and conjugation
CC may be via a bigeneric spacer. The SPNE polypeptides bind an HIV broadly

XX
PD 21-NOV-2002.
XX
PF 15-MAY-2002; 2002WO-US015520.
XX
PR 15-MAY-2001; 2001US-0291118P.
PR 31-OCT-2001; 2001US-00001469.
PR 14-DEC-2001; 2001US-00017666.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Faris M, Raitano AB, Morrison RK, Saffran D, Ge W;
PI Challita-Eid PM;
XX
DR WPI; 2003-129310/12.
XX
PT New composition comprising 101P3A11-related protein, useful for
PT preventing or treating cancer e.g., stomach, cervix, uterus, rectum,
PT prostate, colon, kidney or breast cancer.
XX
PS Claim 13; SEQ ID NO 3890; 327pp; English.
XX
CC The invention relates to a novel composition comprising: a substance that
CC modulates the status of a protein comprising a sequence with a fully
CC defined 2466 or 3136 amino acid sequence given in the specification; or a
CC molecule that is immobilised by a protein comprising a sequence with a
CC fully defined 2466 or 3136 amino acids, where the status of a cell
CC expressing the protein is modulated. The invention further relates to: an
CC antibody; a non-human transgenic animal or hybridoma that produces the
CC antibody; a method of delivering a cytotoxic or diagnostic agent to a
CC cell that expresses the protein; a polynucleotide that encodes the
CC protein; a method for inhibiting the growth of cancer cells that express
CC the protein; a method for generating a mammalian immune response directed
CC to the protein; a method detecting in a sample the presence of a 101P3A11
CC -related protein or polynucleotide; and a method for monitoring one or
CC more 101P3A11 gene products in a biological sample from a patient having
CC or suspected of having cancer. The novel composition has cytostatic
CC activity. The composition is useful for preventing or treating cancer
CC e.g., stomach, cervix, uterus, rectum, prostate, colon, kidney or breast
CC cancer. This sequence represents a peptide derived from the 101P3A11
CC protein of the invention.
XX
SQ Sequence 15 AA;

Query Match 37.7%; Score 29; DB 7; Length 15;
Best Local Similarity 54.5%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 RLPFVKAHVRE 14
||| | | |
Db 5 RLFHVATHASE 15

RESULT 9
AAU77504
ID AAU77504 standard; peptide; 10 AA.
XX
AC AAU77504;
XX
DT 05-JUN-2002 (first entry)
XX
DE Dictyostelium dipeptidylaminopeptidase decapeptide substrate.
XX
KW Dictyostelium dipeptidylaminopeptidase; dDAP.
XX
OS Synthetic.
XX
PN EP1094118-A1.
XX
PD 25-APR-2001.
XX
PF 29-SEP-1993; 2001EP-00200113.
XX
PR 01-OCT-1992; 92US-00955539.
PR 29-SEP-1993; 93EP-00307746.
XX
PA (ELIL) LILLY & CO ELI.
XX
PI Atkinson PR, Hilton MD, Lambooy PK;
XX
DR WPI; 2002-228903/29.
XX
PT Removing amino-terminal dipeptides from a precursor polypeptide, involves
PT contacting the precursor polypeptide with Dictyostelium
PT dipeptidylaminopeptidase enzyme obtainable from Dictyostelium discoideum.
XX
PS Example 3; Page 7; 10pp; English.
XX
CC The invention describes a method of removing an amino-terminal dipeptide
CC from a precursor polypeptide. The method involves contacting the
CC precursor polypeptide with a Dictyostelium dipeptidylaminopeptidase
CC (dDAP) enzyme obtainable from Dictyostelium discoideum. This sequence
CC represents a peptide substrate used to study the action of dDAP
XX
SQ Sequence 10 AA;

Query Match 36.4%; Score 28; DB 5; Length 10;
Best Local Similarity 44.4%; Pred. No. 3.9e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RLFFVKAHV 12
|::|| |:
Db 2 RMYFVNQHL 10

RESULT 10
ABP46458
ID ABP46458 standard; peptide; 14 AA.
XX
AC ABP46458;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human BLyS binding scFv VH CDR3 SEQ ID 2469.
XX
KW BLys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
KW tumour necrosis factor; B cell proliferation; B cell differentiation;
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
KW common variable immunodeficiency; acquired immunodeficiency syndrome.
XX
OS Homo sapiens.
XX
PN WO200202641-A1.
XX
PD 10-JAN-2002.
XX
PF 15-JUN-2001; 2001WO-US019110.
XX
PR 16-JUN-2000; 2000US-0212210P.
PR 17-OCT-2000; 2000US-0240816P.
PR 16-MAR-2001; 2001US-0276248P.
PR 21-MAR-2001; 2001US-0277379P.
PR 25-MAY-2001; 2001US-0293499P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.
XX
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;
XX
DR WPI; 2002-114799/15.
XX
PT Antibodies against B Lymphocyte Stimulating polypeptides, useful for the
PT diagnosis and treatment of cancers and immune disorders.
XX

DE 101P3A11 protein derived peptide, SEQ ID No 3586.
XX
KW transgenic animal; cytotoxic; cancer; immune; 101P3A11; cytostatic;
KW stomach; cervix; uterus; rectum; prostate; colon; kidney; breast.
XX
OS Unidentified.
XX
PN WO200292842-A2.
XX
PD 21-NOV-2002.
XX
PF 15-MAY-2002; 2002WO-US015520.
XX
PR 15-MAY-2001; 2001US-0291118P.
PR 31-OCT-2001; 2001US-00001469.
PR 14-DEC-2001; 2001US-00017666.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Faris M, Raitano AB, Morrison RK, Saffran D, Ge W;
PI Challita-Eid PM;
XX
DR WPI; 2003-129310/12.
XX
PT New composition comprising 101P3A11-related protein, useful for
PT preventing or treating cancer e.g., stomach, cervix, uterus, rectum,
PT prostate, colon, kidney or breast cancer.
XX
PS Claim 13; SEQ ID NO 3586; 327pp; English.
XX
CC The invention relates to a novel composition comprising: a substance that
CC modulates the status of a protein comprising a sequence with a fully
CC defined 2466 or 3136 amino acid sequence given in the specification; or a
CC molecule that is immobilised by a protein comprising a sequence with a
CC fully defined 2466 or 3136 amino acids, where the status of a cell
CC expressing the protein is modulated. The invention further relates to: an
CC antibody; a non-human transgenic animal or hybridoma that produces the
CC antibody; a method of delivering a cytotoxic or diagnostic agent to a
CC cell that expresses the protein; a polynucleotide that encodes the
CC protein; a method for inhibiting the growth of cancer cells that express
CC the protein; a method for generating a mammalian immune response directed
CC to the protein; a method detecting in a sample the presence of a 101P3A11
CC -related protein or polynucleotide; and a method for monitoring one or
CC more 101P3A11 gene products in a biological sample from a patient having
CC or suspected of having cancer. The novel composition has cytostatic
CC activity. The composition is useful for preventing or treating cancer
CC e.g., stomach, cervix, uterus, rectum, prostate, colon, kidney or breast
CC cancer. This sequence represents a peptide derived from the 101P3A11
CC protein of the invention.
XX
SQ Sequence 15 AA;

Query Match 37.7%; Score 29; DB 7; Length 15;
Best Local Similarity 54.5%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 RLFFVKAHVRE 14
||| | | |
Db 4 RLFHVATHASE 14

RESULT 7
ADI09519
ID ADI09519 standard; peptide; 15 AA.
XX
AC ADI09519;
XX
DT 22-APR-2004 (first entry)
XX
DE 101P3A11 protein derived peptide, SEQ ID No 3705.
XX
KW transgenic animal; cytotoxic; cancer; immune; 101P3A11; cytostatic;
KW stomach; cervix; uterus; rectum; prostate; colon; kidney; breast.
XX
OS Unidentified.
XX
PN WO200292842-A2.
XX
PD 21-NOV-2002.
XX
PF 15-MAY-2002; 2002WO-US015520.
XX
PR 15-MAY-2001; 2001US-0291118P.
PR 31-OCT-2001; 2001US-00001469.
PR 14-DEC-2001; 2001US-00017666.
XX
PA (AGEN-) AGENSYS INC.
XX
PI Jakobovits A, Faris M, Raitano AB, Morrison RK, Saffran D, Ge W;
PI Challita-Eid PM;
XX
DR WPI; 2003-129310/12.
XX
PT New composition comprising 101P3A11-related protein, useful for
PT preventing or treating cancer e.g., stomach, cervix, uterus, rectum,
PT prostate, colon, kidney or breast cancer.
XX
PS Claim 13; SEQ ID NO 3705; 327pp; English.
XX
CC The invention relates to a novel composition comprising: a substance that
CC modulates the status of a protein comprising a sequence with a fully
CC defined 2466 or 3136 amino acid sequence given in the specification; or a
CC molecule that is immobilised by a protein comprising a sequence with a
CC fully defined 2466 or 3136 amino acids, where the status of a cell
CC expressing the protein is modulated. The invention further relates to: an
CC antibody; a non-human transgenic animal or hybridoma that produces the
CC antibody; a method of delivering a cytotoxic or diagnostic agent to a
CC cell that expresses the protein; a polynucleotide that encodes the
CC protein; a method for inhibiting the growth of cancer cells that express
CC the protein; a method for generating a mammalian immune response directed
CC to the protein; a method detecting in a sample the presence of a 101P3A11
CC -related protein or polynucleotide; and a method for monitoring one or
CC more 101P3A11 gene products in a biological sample from a patient having
CC or suspected of having cancer. The novel composition has cytostatic
CC activity. The composition is useful for preventing or treating cancer
CC e.g., stomach, cervix, uterus, rectum, prostate, colon, kidney or breast
CC cancer. This sequence represents a peptide derived from the 101P3A11
CC protein of the invention.
XX
SQ Sequence 15 AA;

Query Match 37.7%; Score 29; DB 7; Length 15;
Best Local Similarity 54.5%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 4 RLFFVKAHVRE 14
||| | | |
Db 5 RLFHVATHASE 15

RESULT 8
ADI09704
ID ADI09704 standard; peptide; 15 AA.
XX
AC ADI09704;
XX
DT 22-APR-2004 (first entry)
XX
DE 101P3A11 protein derived peptide, SEQ ID No 3890.
XX
KW transgenic animal; cytotoxic; cancer; immune; 101P3A11; cytostatic;
KW stomach; cervix; uterus; rectum; prostate; colon; kidney; breast.
XX
OS Unidentified.
XX
PN WO200292842-A2.

```
SQ   Sequence 14 AA;

  Query Match             40.3%;  Score 31;  DB 4;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 GRRLFFVKAHVRES 15
              |  |  :  ||:||
Db          1 GAMLLNISGHVKES 14


RESULT 4
ABG76689
ID   ABG76689 standard; peptide; 15 AA.
XX
AC   ABG76689;
XX
DT   05-NOV-2002  (first entry)
XX
DE   Hepatitis C virus (HCV)-specific ligand #20.
XX
KW   Binding specificity; anti-antigen antibody; serum; ADAM;
KW   antigen detection by antigen mimic; Hepatitis C virus infection; HCV;
KW   HCV-specific ligand; human.
XX
OS   Homo sapiens.
XX
PN   WO200237115-A1.
XX
PD   10-MAY-2002.
XX
PF   03-NOV-2000; 2000WO-IT000442.
XX
PR   03-NOV-2000; 2000WO-IT000442.
XX
PA   (KENT-) KENTON SRL.
XX
PI   Felici F,  Gargano N,  Minenkova O,  Monaci P;
XX
DR   WPI; 2002-599299/64.
XX
PT   Making diagnosis of antigen, by identifying binding specificity of anti-
PT   antigen antibody molecules in serum by antibody detection by antigen
PT   mimics methodology, and identifying antibodies associated with antigen.
XX
PS   Claim 26; Page 42; 86pp; English.
XX
CC   The present invention relates to a method for making a diagnosis of an
CC   antigen. The method involves identifying the binding specificity of the
CC   anti-antigen antibody molecules in serum by the antibody detection by
CC   antigen mimics (ADAM) methodology. The method comprises screening phage
CC   libraries using sera from antigen-infected and non-infected individuals,
CC   and identifying peptides binding antibodies (ligands) specifically
CC   associated with the antigen. The method of the invention can be used for
CC   the detection of infectious agents, particularly Hepatitis C virus (HCV).
CC   The invention provides HCV-specific ligands which are useful for the
CC   preparation of a diagnostic assay for detecting HCV infection in a
CC   subject. The HCV-specific ligands are also useful for the preparation of
CC   vaccines against HCV. ABG76670-ABG76743 represent HCV-specific ligands
CC   identified from human sera
XX
SQ   Sequence 15 AA;

  Query Match             40.3%;  Score 31;  DB 5;  Length 15;
  Best Local Similarity   58.3%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 RLFFVKAHVRES 15
              | | :|||  ||
Db          3 RTFLMKAHGIES 14


RESULT 5
ABB56587
ID   ABB56587 standard; peptide; 14 AA.
XX
AC   ABB56587;
XX
DT   05-MAR-2002  (first entry)
XX
DE   Human SNP related amino acid sequence SEQ ID NO:1152.
XX
KW   Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;
KW   immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW   autoimmune disease; inflammation; cancer; nervous system disease;
KW   infection; polymorphic protein.
XX
OS   Homo sapiens.
XX
PN   WO200138586-A2.
XX
PD   31-MAY-2001.
XX
PF   22-NOV-2000; 2000WO-US032311.
XX
PR   24-NOV-1999;   99US-0167383P.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-355949/37.
XX
PT   Isolated human nucleic acids comprising one or more single nucleotide
PT   polymorphisms, useful for treating a subject suffering from a pathology,
PT   e.g. autoimmune diseases, ascribed to the presence of a sequence
PT   polymorphism.
XX
PS   Claim 1; Page 597; 674pp; English.
XX
CC   ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC   comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC   to ABB56903 represent human peptides encoded by some of the SNP
CC   oligonucleotides. The sequences from the present invention can have
CC   immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC   antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC   and antibodies from the present invention can be used for treating a
CC   subject suffering from, at risk for, or suspected of, suffering from a
CC   pathology ascribed to the presence of a sequence polymorphism. The
CC   pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC   the nervous system, and infection by pathogenic microorganisms. The SNPs
CC   are also useful for determining which forms of a characterised
CC   polymorphism are present in individuals. The antibodies may be used in
CC   the detection, quantitation and/or cellular or tissue localisation of a
CC   polymorphic protein (e.g., for use in measuring levels of the polymorphic
CC   protein within appropriate physiological samples)
XX
SQ   Sequence 14 AA;

  Query Match             39.0%;  Score 30;  DB 4;  Length 14;
  Best Local Similarity   62.5%;  Pred. No. 2.4e+02;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRRLFFVK 9
              |:||| :|
Db          1 GKRLFVIK 8


RESULT 6
ADI09400
ID   ADI09400 standard; peptide; 15 AA.
XX
AC   ADI09400;
XX
DT   22-APR-2004  (first entry)
XX
```

```
CC   simplex virus DNA polymerase and is implicated as a foreign epitope
CC   involved in the aetiology or in remissions of multiple sclerosis. It has
CC   been shown capable of inducing the proliferation of autoreactive T-cell
CC   clones isolated from multiple sclerosis patients
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 77;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 9.7e-07;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGRRLPFVKAHVRES 15
              |||||||||||||||
Db          1 GGRRLPFVKAHVRES 15


RESULT 2
AAW04852
ID   AAW04852 standard; peptide; 15 AA.
XX
AC   AAW04852;
XX
DT   16-OCT-2003  (revised)
DT   18-FEB-1997  (first entry)
XX
DE   Internal fragment of Epstein-Barr virus DNA polymerase.
XX
KW   Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW   HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW   desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW   phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW   DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS   Human herpesvirus 4.
XX
PN   WO9627387-A1.
XX
PD   12-SEP-1996.
XX
PF   07-MAR-1996;   96WO-US003182.
XX
PR   07-MAR-1995;   95US-00400796.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Strominger JL,  Wucherpfennig KW;
XX
DR   WPI; 1996-425218/42.
XX
PT   Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT   - useful in disease treatment, and method for identification of other
PT   self and non-self antigens implicated in auto-immune disease.
XX
PS   Claim 2; Page 45; 58pp; English.
XX
CC   Pharmaceutical preparations for tolerisation to antigens comprise either
CC   an isolated human non-collagen or non-myslin basic protein (MBP)
CC   polypeptide which is capable of tolerising an individual to an
CC   autoantigen; or an isolated human pathogen polypeptide capable of
CC   tolerising an individual to that polypeptide. In both cases, the
CC   polypeptide (whether self or non-self) includes an amino acid sequence
CC   corresponding to a sequence motif for a MHC class II protein, such as HLA
CC   -DR, which is associated with a human autoimmune disease and which binds
CC   to the polypeptide to activate autoreactive T-cells in individuals with
CC   the autoimmune disease. This peptide is an internal peptide of Epstein-
CC   Barr virus DNA polymerase and is implicated as a foreign epitope involved
CC   in the aetiology or in remissions of multiple sclerosis. It has been
CC   shown capable of inducing the proliferation of autoreactive T-cell clones
CC   isolated from multiple sclerosis patients. (Updated on 16-OCT-2003 to
CC   standardise OS field)
XX
SQ   Sequence 15 AA;

  Query Match             45.5%;  Score 35;  DB 2;  Length 15;
  Best Local Similarity   77.8%;  Pred. No. 33;
  Matches    7;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          7 FVKAHVRES 15
              ||| || ||
Db          7 FVKKHVHES 15


RESULT 3
AAG98020
ID   AAG98020 standard; peptide; 14 AA.
XX
AC   AAG98020;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human SNP associated peptide SEQ ID NO. 662.
XX
KW   Human; single nucleotide polymorphism; SNP; angiopoietin;
KW   4-hydroxybutyrate; dehydrogenase; protein therapy;
KW   adenosine triphosphate-dependent RNA helicase;
KW   major histocompatibility complex Class I histocompatibility antigen; MHC;
KW   phosphoglycerate kinase; immunosuppressive; immunostimulatory;
KW   antirheumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;
KW   antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine.
XX
OS   Homo sapiens.
XX
PN   WO200148245-A2.
XX
PD   05-JUL-2001.
XX
PF   27-DEC-2000; 2000WO-US035346.
XX
PR   27-DEC-1999;   99US-00472688.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-418297/44.
XX
PT   Polymorphic nucleic acids encoding e.g. angiopoietin, dehydrogenase,
PT   adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate
PT   kinase, useful for diagnosing and treating, e.g. cancer, autoimmune
PT   diseases and infections.
XX
PS   Disclosure; Page 436; 484pp; English.
XX
CC   The invention relates to nucleic acids (AAH79386-AAH80036) encoding
CC   polymorphic variants of proteins (AAG98010-AAG98238) related to
CC   angiopoietin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate
CC   (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)
CC   Class I histocompatibility antigen and/or phosphoglycerate kinase. These
CC   nucleic acid single nucleotide polymorphisms (SNPs) and the encoded
CC   proteins have potential immunosuppressive, immunostimulatory,
CC   antirheumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,
CC   antileukemic, neuroprotective and antimicrobial activity and may be
CC   useful in gene/protein therapy, vaccines, modulation of the expression
CC   and activity of proteins related to angiopoietin, 4-hydroxybutyrate,
CC   dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase, major
CC   histocompatibility complex (MHC) Class I histocompatibility antigen
CC   and/or phosphoglycerate kinase. Disorders that may be prevented,
CC   diagnosed and/or treated by the above methods include multifactorial
CC   diseases with a genetic component, such as autoimmune diseases (e.g.
CC   rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus
CC   erythematosus and Grave's disease), inflammation, cancer (e.g. cancers of
CC   the bladder, brain, breast, colon and kidney, leukemia), diseases of the
CC   nervous system, an infection of pathogenic organisms. They may also be
CC   used to alter phenotypic traits such as longevity, appearance, strength,
CC   speed and endurance
XX
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-15
Perfect score: 77
Sequence: 1 GGRRLFFVKAHVRES 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	77	100.0	15	2	AAW04855	Aaw04855 Internal
2	35	45.5	15	2	AAW04852	Aaw04852 Internal
3	31	40.3	14	4	AAG98020	Aag98020 Human SNP
4	31	40.3	15	5	ABG76689	Abg76689 Hepatitis
5	30	39.0	14	4	ABB56587	Abb56587 Human SNP
6	29	37.7	15	7	ADI09400	Adi09400 101P3A11
7	29	37.7	15	7	ADI09519	Adi09519 101P3A11
8	29	37.7	15	7	ADI09704	Adi09704 101P3A11
9	28	36.4	10	5	AAU77504	Aau77504 Dictyoste
10	28	36.4	14	5	ABP46458	Abp46458 Human BLy
11	28	36.4	14	7	ADG97285	Adg97285 scFV VHCD
12	28	36.4	15	2	AAR93674	Aar93674 HIV princ
13	28	36.4	15	5	ABG76687	Abg76687 Hepatitis
14	27	35.1	8	4	AAJ02281	Aaj02281 Hepatitis
15	27	35.1	9	4	AAJ01219	Aaj01219 Hepatitis
16	27	35.1	9	4	AAJ02683	Aaj02683 Hepatitis
17	27	35.1	9	4	AAJ02187	Aaj02187 Hepatitis
18	27	35.1	10	4	AAJ02282	Aaj02282 Hepatitis
19	27	35.1	10	4	AAJ02744	Aaj02744 Hepatitis
20	27	35.1	11	4	AAJ02684	Aaj02684 Hepatitis
21	27	35.1	11	4	AAJ02188	Aaj02188 Hepatitis
22	27	35.1	11	6	ABP76436	Abp76436 Peptidomi
23	27	35.1	12	8	ADJ34590	Adj34590 AKT subst
24	27	35.1	13	4	AAG84591	Aag84591 MAGE3 DR
25	27	35.1	13	4	AAG84606	Aag84606 MAGE3 DR
26	27	35.1	13	4	AAG84613	Aag84613 MAGE3 DR
27	27	35.1	13	4	AAG84618	Aag84618 MAGE3 DR
28	27	35.1	14	4	AAG98022	Aag98022 Human SNP
29	27	35.1	15	5	ABG76686	Abg76686 Hepatitis
30	27	35.1	15	7	ADE86173	Ade86173 Antennape
31	27	35.1	15	7	ADE78445	Ade78445 Peptide t
32	27	35.1	15	8	ADI35013	Adi35013 Hepatitis
33	27	35.1	15	8	ADI35012	Adi35012 Hepatitis
34	27	35.1	15	8	ADL26316	Adl26316 Synthetic
35	26	33.8	7	2	AAW10790	Aaw10790 Ferritin
36	26	33.8	7	2	AAW10763	Aaw10763 Ferritin
37	26	33.8	7	2	AAW10808	Aaw10808 Ferritin
38	26	33.8	7	5	ABG77576	Abg77576 Targettin
39	26	33.8	8	2	AAW31837	Aaw31837 Peptide 3
40	26	33.8	8	4	AAU05710	Aau05710 p21 C-ter
41	26	33.8	9	2	AAR89238	Aar89238 LN clone
42	26	33.8	9	4	AAE11956	Aae11956 Mammalian
43	26	33.8	9	6	ABP76390	Abp76390 Peptidomi
44	26	33.8	9	8	ADK04523	Adk04523 Hepatitis
45	26	33.8	9	8	ADK04077	Adk04077 Hepatitis

ALIGNMENTS

RESULT 1
AAW04855
ID AAW04855 standard; peptide; 15 AA.
XX
AC AAW04855;
XX
DT 18-FEB-1997 (first entry)
XX
DE Internal fragment of herpes simplex virus DNA polymerase.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Herpes simplex virus.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 2; Page 47; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with
CC the autoimmune disease. This peptide is an internal peptide of herpes

```
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Moniliformopses; Filicophyta; Marattiopsida; Marattiales;
OC   Marattiaceae; Archangiopteris.
OX   NCBI_TaxID=203826;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Chiang T., Chiang Y., Chou C., Cheng Y., Chiou W.;
RT   "Phylogeography and conservation of Archangiopteris somai and A. itoi
RT   (Marattiaceae, Pteridophyta) based on nucleotide variation of cpDNA
RT   atpB-rbcL intergenic spacer.";
RL   Submitted (AUG-2004) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Chiang Y.C.;
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ505259; CAD44048.1; -.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
KW   Chloroplast; Hydrolase.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1195 MW;  9E2AF0C9C7376451 CRC64;

  Query Match             26.0%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1e+04;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LFFV 8
              ||||
Db          6 LFFV 9


RESULT 15
O39952
ID   O39952      PRELIMINARY;      PRT;    10 AA.
AC   O39952;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   E1 protein (Fragment).
OS   GB virus C/Hepatitis G virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   GBV-C/HGV group.
OX   NCBI_TaxID=54290;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97368412; PubMed=9225026;
RA   Smith D.B., Cuceanu N., Davidson F., Jarvis L.M., Mokili J.L.,
RA   Hamid S., Ludlam C.A., Simmonds P.;
RT   "Discrimination of hepatitis G virus/GBV-C geographical variants by
RT   analysis of the 5' non-coding region.";
RL   J. Gen. Virol. 78:1533-1542(1997).
DR   EMBL; AF003170; AAC57981.1; -.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1152 MW;  CC88F0C9C7272732 CRC64;

  Query Match             26.0%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 1e+04;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 LFFV 8
              ||||
Db          6 LFFV 9


Search completed: February 22, 2005, 09:38:06
Job time : 53.6667 secs
```

```
SQ   SEQUENCE   11 AA;  1370 MW;  592BC02D12C9DB57 CRC64;

  Query Match             27.3%;  Score 21;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 7.1e+03;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 RLFFVK 9
              :|:||:
Db          3 KLYFVR 8


RESULT 11
G6PD_MAIZE
ID   G6PD_MAIZE     STANDARD;      PRT;    15 AA.
AC   P80619;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Glucose-6-phosphate 1-dehydrogenase, cytoplasmic isoform (EC 1.1.1.49)
DE   (G6PD) (2D-page of etiolated coleoptile spot 243) (Fragment).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX   NCBI_TaxID=4577;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Coleoptile;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- CATALYTIC ACTIVITY: D-glucose 6-phosphate + NADP(+) = D-glucono-
CC       1,5-lactone 6-phosphate + NADPH.
CC   -!- PATHWAY: Pentose phosphate pathway; first step.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the glucose-6-phosphate dehydrogenase
CC       family.
DR   Maize-2DPAGE; P80619; COLEOPTILE.
DR   MaizeDB; 123946; -.
DR   InterPro; IPR001282; G6PD.
DR   PROSITE; PS00069; G6P_DEHYDROGENASE; PARTIAL.
KW   Carbohydrate metabolism; Direct protein sequencing;
KW   Glucose metabolism; NADP; Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1739 MW;  02038EE7471AE038 CRC64;

  Query Match             27.3%;  Score 21;  DB 1;  Length 15;
  Best Local Similarity   28.6%;  Pred. No. 9.7e+03;
  Matches    4;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 GRRLFFVKAHVRES 15
              ||  | ::    |:
Db          2 GRNEFVIRLQXSEA 15


RESULT 12
O84887
ID   O84887      PRELIMINARY;      PRT;    15 AA.
AC   O84887;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Truncated GTP-binding protein.
GN   Name=lepA;
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RA   Figueroa-Bossi N., Bossi L.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF001386; AAC26066.1; -.
DR   PIR; T03000; T03000.
SQ   SEQUENCE   15 AA;  1786 MW;  9DC811919C7EA87C CRC64;

  Query Match             27.3%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 9.7e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 RLFFVKAHV 12
              | | : ||:
Db          5 RNFSIIAHI 13


RESULT 13
TPIS_NICPL
ID   TPIS_NICPL     STANDARD;      PRT;    10 AA.
AC   P19118;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Triosephosphate isomerase, cytosolic (EC 5.3.1.1) (TIM) (Triose-
DE   phosphate isomerase) (Fragment).
OS   Nicotiana plumbaginifolia (Leadwort-leaved tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4092;
RN   [1]
RP   SEQUENCE.
RA   Bauw G., de Loose M., Inze D., van Montagu M., Vandekerckhove J.;
RT   "Alterations in the phenotype of plant cells studied by NH2-terminal
RT   amino acid-sequence analysis of proteins electroblotted from two-
RT   dimensional gel-separated total extracts.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:4806-4810(1987).
CC   -!- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glycerone
CC       phosphate.
CC   -!- PATHWAY: Plays an important role in several metabolic pathways.
CC   -!- SUBUNIT: Homodimer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC   -!- MISCELLANEOUS: In plants, there are two types of TPIS, cytosolic
CC       and plastid.
CC   -!- SIMILARITY: Belongs to the triosephosphate isomerase family.
DR   PIR; A27617; A27617.
DR   InterPro; IPR000652; Triophos_ismrse.
DR   PROSITE; PS00171; TIM; PARTIAL.
KW   Direct protein sequencing; Fatty acid biosynthesis; Gluconeogenesis;
KW   Glycolysis; Isomerase; Pentose shunt.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1140 MW;  80B9D37862C9C9D1 CRC64;

  Query Match             26.0%;  Score 20;  DB 1;  Length 10;
  Best Local Similarity   80.0%;  Pred. No. 1e+04;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 RLFFV 8
              | |||
Db          2 RTFFV 6


RESULT 14
Q6A3T0
ID   Q6A3T0      PRELIMINARY;      PRT;    10 AA.
AC   Q6A3T0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   ATP synthase beta subunit (EC 3.6.3.14) (Fragment).
GN   Name=atpB;
OS   Archangiopteris somai.
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=95079448; PubMed=7527301;
RA   Okada H., Yoshida J., Seo H., Wakabayashi T., Sugita K., Hagiwara M.;
RT   "Anti-(glioma surface antigen) monoclonal antibody G-22 recognizes
RT   overexpressed CD44 in glioma cells.";
RL   Cancer Immunol. Immunother. 39:313-317(1994).
FT   NON_TER       1      1
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1337 MW;  2E0F6CE9D9D2C1E8 CRC64;

  Query Match             28.6%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 4.9e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 FVKAHV 12
              |:: ||
Db          2 FIEGHV 7


RESULT 7
Q7M2Q3
ID   Q7M2Q3      PRELIMINARY;      PRT;    14 AA.
AC   Q7M2Q3;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Probursin tetradecapeptide.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=91133172; PubMed=1671611; DOI=10.1016/0024-3205(91)90092-P;
RA   Audhya T., King R., Goldstein G.;
RT   "Bovine probursin tetradecapeptide contains amino acid sequence from
RT   somatostatin, tuftsin and bursin.";
RL   Life Sci. 48:773-780(1991).
DR   PIR; JH0328; JH0328.
SQ   SEQUENCE   14 AA;  1801 MW;  2CF3549CEC6BFA17 CRC64;

  Query Match             28.6%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 5.8e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGRR 4
              ||||
Db         11 GGRR 14


RESULT 8
Q95949
ID   Q95949      PRELIMINARY;      PRT;    15 AA.
AC   Q95949;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   NADH dehydrogenase subunit 4 (Fragment).
GN   Name=ND4;
OS   Sauromalus ater (Common chuckwalla) (Sauromalus obesus).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Iguania; Iguanidae; Iguaninae; Sauromalus.
OX   NCBI_TaxID=65995;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97019047; PubMed=8865663;
RA   Sites J.W. Jr., Davis S.K., Guerra T., Iverson J.B., Snell H.L.;
RT   "Character congruence and phylogenetic signal in molecular and
RT   morphological data sets: a case study in the living Iguanas (Squamata,
RT   Iguanidae).";
RL   Mol. Biol. Evol. 13:1087-1105(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sites J.W., Davis S.K., Guerra T., Iverson J.B., Snell H.L.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U66232; AAB07479.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1896 MW;  8327179CFC353901 CRC64;

  Query Match             28.6%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   30.8%;  Pred. No. 6.2e+03;
  Matches    4;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 RRLFFVKAHVRES 15
              |||   : |: ::
Db          3 RRLHSTRRHLTKT 15


RESULT 9
Q37854
ID   Q37854      PRELIMINARY;      PRT;     8 AA.
AC   Q37854;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Coliphage gene of unknown function, 5'end. (Fragment).
OS   Bacteriophage R17.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Leviviridae;
OC   Levivirus.
OX   NCBI_TaxID=12026;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=73224987; PubMed=4352721;
RA   Rensing U.F.E.;
RT   "A sequence of seventy-three nucleotides from the coliphage R17
RT   genome.";
RL   Biochem. J. 131:593-604(1973).
DR   EMBL; M24820; AAA72755.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  969 MW;  ECB45412C1E72726 CRC64;

  Query Match             27.3%;  Score 21;  DB 2;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 HVRES 15
              ||| |
Db          4 HVRNS 8


RESULT 10
Q6RJX9
ID   Q6RJX9      PRELIMINARY;      PRT;    11 AA.
AC   Q6RJX9;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Induced stolon tip protein PJ-2.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC   lamiids; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kim S., Lee K.-W.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY496132; AAR83870.1; -.
```

```
RESULT 2
Q6QQR3
ID   Q6QQR3      PRELIMINARY;      PRT;    15 AA.
AC   Q6QQR3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=HP0218;
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Helicobacteraceae; Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=92-28;
RA   Terry C., Madigan K., McGinnis L., Cao P., Cover T., Liechti G.,
RA   Peek R. Jr., Forsyth M.;
RL   Submitted (JAN-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY529682; AAS20951.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1883 MW;  59AF5B291B125652 CRC64;

  Query Match             36.4%;  Score 28;  DB 2;  Length 15;
  Best Local Similarity   46.2%;  Pred. No. 4.3e+02;
  Matches    6;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 GRRLFFVKAHVRE 14
              ||: |  |  ||:
Db          3 GRKEFLYKQFVRK 15


RESULT 3
Q945F2
ID   Q945F2      PRELIMINARY;      PRT;    14 AA.
AC   Q945F2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Aldolase (Fragment).
OS   Cicer arietinum (Chickpea) (Garbanzo).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Cicereae; Cicer.
OX   NCBI_TaxID=3827;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Rajesh P.N., Gupta V.S., Ranjekar P.K., Muehlbauer F.J.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF416480; AAL08017.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1785 MW;  421376873F9353C7 CRC64;

  Query Match             31.2%;  Score 24;  DB 2;  Length 14;
  Best Local Similarity   55.6%;  Pred. No. 2.4e+03;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 RRLFFVKAH 11
              ||  ||| :
Db          4 RRELFVKGY 12


RESULT 4
Q79CE7
ID   Q79CE7      PRELIMINARY;      PRT;    11 AA.
AC   Q79CE7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   CtrA protein (Fragment).
GN   Name=ctrA;
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC   Neisseriaceae; Neisseria.
OX   NCBI_TaxID=487;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMB;
RX   MEDLINE=96312345; PubMed=8763931;
RA   Swartley J.S., Ahn J.H., Liu L.J., Kahler C.M., Stephens D.S.;
RT   "Expression of sialic acid and polysialic acid in serogroup B
RT   Neisseria meningitidis: divergent transcription of biosynthesis and
RT   transport operons through a common promoter region.";
RL   J. Bacteriol. 178:4052-4059(1996).
DR   EMBL; U40741; AAB17456.1; -.
FT   NON_TER       11     11
SQ   SEQUENCE   11 AA;  1440 MW;  DB2075394B59C322 CRC64;

  Query Match             29.9%;  Score 23;  DB 2;  Length 11;
  Best Local Similarity   36.4%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 LFFVKAHVRES 15
              :| || ::| :
Db          1 MFKVKFYIRHA 11


RESULT 5
Q9X9S6
ID   Q9X9S6      PRELIMINARY;      PRT;    11 AA.
AC   Q9X9S6;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein ORF9 (Fragment).
GN   Name=ORF9;
OS   Streptomyces lividans.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1916;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TK21;
RX   MEDLINE=99328982; PubMed=10400594;
RA   Martinez-Costa O.H., Martin-Triana A.J., Martinez E.,
RA   Fernandez-Moreno M.A., Malpartida F.;
RT   "An additinal regulatory gene for actinorhodin production in
RT   Streptomyces lividans involves a LysR-type transcriptional
RT   regulator.";
RL   J. Bacteriol. 181:4353-4364(1999).
DR   EMBL; Y18818; CAB51138.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   11 AA;  1160 MW;  D1BABA8EC1EDC412 CRC64;

  Query Match             28.6%;  Score 22;  DB 2;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 4.5e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 VKAHV 12
              |:|||
Db          4 VRAHV 8


RESULT 6
Q9UC29
ID   Q9UC29      PRELIMINARY;      PRT;    12 AA.
AC   Q9UC29;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   85 kDa glioma membrane protein/CD44 homolog (Fragment).
OS   Homo sapiens (Human).
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-15
Perfect score: 77
Sequence: 1 GGRRLFFVKAHVRES 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	37.7	12	1	YZPY_ECOLI	P17776 escherichia
2	28	36.4	15	2	Q6QQR3	Q6qqr3 helicobacte
3	24	31.2	14	2	Q945F2	Q945f2 cicer ariet
4	23	29.9	11	2	Q79CE7	Q79ce7 neisseria m
5	22	28.6	11	2	Q9X9S6	Q9x9s6 streptomyce
6	22	28.6	12	2	Q9UC29	Q9uc29 homo sapien
7	22	28.6	14	2	Q7M2Q3	Q7m2q3 bos taurus
8	22	28.6	15	2	Q95949	Q95949 sauromalus
9	21	27.3	8	2	Q37854	Q37854 bacteriopha
10	21	27.3	11	2	Q6RJX9	Q6rjx9 capsicum an
11	21	27.3	15	1	G6PD_MAIZE	P80619 zea mays (m
12	21	27.3	15	2	O84887	O84887 salmonella
13	20	26.0	10	1	TPIS_NICPL	P19118 nicotiana p
14	20	26.0	10	2	Q6A3T0	Q6a3t0 archangiopt
15	20	26.0	10	2	O39952	O39952 gb virus c/
16	20	26.0	10	2	Q9WLE4	Q9wle4 gb virus c/
17	20	26.0	11	2	Q8MER8	Q8mer8 dombeya til
18	20	26.0	11	2	Q9RBV0	Q9rbv0 pseudomonas
19	20	26.0	11	2	O39951	O39951 gb virus c/
20	20	26.0	14	2	Q96Q50	Q96q50 homo sapien
21	20	26.0	15	1	RM12_YEAST	P36522 saccharomyc
22	20	26.0	15	2	Q9TR46	Q9tr46 bos taurus
23	19	24.7	8	2	Q6EGT1	Q6egt1 stemphylium
24	19	24.7	8	2	Q6EGV0	Q6egv0 stemphylium
25	19	24.7	8	2	Q6EGV3	Q6egv3 stemphylium
26	19	24.7	8	2	Q6EGW2	Q6egw2 stemphylium
27	19	24.7	10	2	Q7M4X7	Q7m4x7 fusarium sp
28	19	24.7	11	2	Q7KYZ8	Q7kyz8 homo sapien
29	19	24.7	12	2	Q7RHA4	Q7rha4 plasmodium
30	19	24.7	12	2	Q945C3	Q945c3 crypthecodi
31	19	24.7	14	1	SAP2_ARBPU	P11760 arbacia pun
32	19	24.7	14	2	Q61864	Q61864 mus musculu
33	19	24.7	15	2	Q9TWF3	Q9twf3 dictyosteli
34	19	24.7	15	2	Q842L7	Q84217 salmonella
35	18	23.4	9	2	Q8MEM3	Q8mem3 howittia tr
36	18	23.4	9	2	Q95GN1	Q95gn1 pelargonium
37	18	23.4	10	2	Q7SA62	Q7sa62 neurospora
38	18	23.4	10	2	Q71UL3	Q71ul3 haliotis ru
39	18	23.4	10	2	Q8HUB4	Q8hub4 anomobryum
40	18	23.4	10	2	Q67B22	Q67b22 bacteriopha
41	18	23.4	10	2	Q7LZJ7	Q7lzj7 carcharhinu
42	18	23.4	11	1	ASL2_BACSE	P83147 bacteroides
43	18	23.4	11	1	MLG_THETS	P41989 theromyzon
44	18	23.4	11	2	O77880	O77880 oreochromis
45	18	23.4	11	2	O77907	O77907 oreochromis

ALIGNMENTS

RESULT 1
YZPY_ECOLI
ID YZPY_ECOLI STANDARD; PRT; 12 AA.
AC P17776;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hypothetical pyrE leader peptide.
GN Name=pyrL; Synonyms=pyrE-LP;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85003588; PubMed=6207018;
RA Poulsen P., Bonekamp F., Jensen K.F.;
RT "Structure of the Escherichia coli pyrE operon and control of pyrE
RT expression by a UTP modulated intercistronic attentuation.";
RL EMBO J. 3:1783-1790(1984).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=83287414; PubMed=6349999;
RA Poulsen P., Jensen K.F., Valentin-Hansen P., Carlsson P.,
RA Lundberg L.G.;
RT "Nucleotide sequence of the Escherichia coli pyrE gene and of the DNA
RT in front of the protein-coding region.";
RL Eur. J. Biochem. 135:223-229(1983).
CC -!- CAUTION: This sequence, according to the EcoSeq database (K. Rudd)
CC is probably not a real protein; therefore this entry will probably
CC be deleted in future releases.
CC ---

CC ---
DR EMBL; X00781; -; NOT_ANNOTATED_CDS.
DR EMBL; V01578; -; NOT_ANNOTATED_CDS.
DR PIR; A30400; LFECPE.
KW Hypothetical protein.
SQ SEQUENCE 12 AA; 1542 MW; C4291FA437A2C9C9 CRC64;

Query Match 37.7%; Score 29; DB 1; Length 12;
Best Local Similarity 70.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          4 RLFFVKAHVR 13
              |||||   ||
Db          3 RLFFVCRKVR 12
```

```
  Query Match             26.0%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 2.7e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRRLFF 7
              |: :||
Db          6 GQEVFF 11


RESULT 14
PH1466
T-cell receptor beta chain (clone A3/74.1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 15-Mar-2004
C;Accession: PH1466
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Ko
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompatib
A;Reference number: PH1430; MUID:93171821; PMID:8436911
A;Accession: PH1466
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Keywords: receptor; T-cell

  Query Match             26.0%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 2.7e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRRLFF 7
              |: :||
Db          6 GQEVFF 11


RESULT 15
S22761
Ig lambda-2 chain J region - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 16-Aug-1996
C;Accession: S22761
R;Weiss, S.; Wu, G.E.
EMBO J. 6, 927-932, 1987
A;Title: Somatic point mutations in unrearranged immunoglobulin gene segments encoding t
A;Reference number: S22759; MUID:87246527; PMID:3109891
A;Accession: S22761
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-13 <WEI>
A;Cross-references: EMBL:X58422
C;Keywords: heterotetramer; immunoglobulin

  Query Match             26.0%;  Score 20;  DB 2;  Length 13;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGRRL 5
              || ||
Db          4 GGTRL 8


Search completed: February 22, 2005, 09:46:30
Job time : 11.1333 secs
```

```
C;Species: Bos primigenius taurus (cattle)
C;Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 16-Aug-2004
C;Accession: JH0328
R;Audhya, T.; King, R.; Goldstein, G.
Life Sci. 48, 773-780, 1991
A;Title: Bovine probursin tetradecapeptide contains amino acid sequence from somatostati
A;Reference number: JH0328; MUID:91133172; PMID:1671611
A;Accession: JH0328
A;Molecule type: protein
A;Residues: 1-14 <AUD>
A;Cross-references: UNIPROT:Q7M2Q3
C;Comment: Intact probursin has the biological activity of both somatostatin and bursin.
C;Keywords: amidated carboxyl end; hormone
F;5-8/Product: tuftsin #status predicted <TUF>
F;9-11/Product: bursin #status predicted <BUR>
F;11/Modified site: amidated carboxyl end (Gly) (amide in mature form from following gly

  Query Match             28.6%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGRR 4
              ||||
Db         11 GGRR 14


RESULT 9
S57577
T cell receptor V-J junctional alpha chain region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57577
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A;Reference number: S57494
A;Accession: S57577
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <BUR>
A;Cross-references: EMBL:Z49945; NID:g887492; PIDN:CAA90216.1; PID:g887493
C;Keywords: T-cell receptor

  Query Match             28.6%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 GGRRLFF 7
              |   |||
Db          8 GANNLFF 14


RESULT 10
T03000
GTP-binding protein, truncated - Salmonella typhimurium
C;Species: Salmonella typhimurium
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03000
R;Figueroa-Bossi, N.; Bossi, L.
submitted to the EMBL Data Library, June 1998
A;Reference number: Z14818
A;Accession: T03000
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-15 <FIG>
A;Cross-references: UNIPROT:O84887; EMBL:AF001386; NID:g3294471; PIDN:AAC26066.1; PID:g3
C;Genetics:
A;Note: lepA

  Query Match             27.3%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 RLFFVKAHV 12
              | | : ||:
Db          5 RNFSIIAHI 13


RESULT 11
A27617
triose-phosphate isomerase (EC 5.3.1.1) - curled-leaved tobacco (fragment)
C;Species: Nicotiana plumbaginifolia (curled-leaved tobacco)
C;Date: 19-May-1989 #sequence_revision 19-May-1989 #text_change 09-Jul-2004
C;Accession: A27617
R;Bauw, G.; De Loose, M.; Inze, D.; Van Montagu, M.; Vandekerckhove, J.
Proc. Natl. Acad. Sci. U.S.A. 84, 4806-4810, 1987
A;Title: Alterations in the phenotype of plant cells studied by NH2-terminal amino acid
A;Reference number: A94167
A;Accession: A27617
A;Molecule type: protein
A;Residues: 1-10 <BAU>
A;Cross-references: UNIPROT:P19118
C;Keywords: gluconeogenesis; glycolysis; intramolecular oxidoreductase; isomerase; pento

  Query Match             26.0%;  Score 20;  DB 2;  Length 10;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 RLFFV 8
              | |||
Db          2 RTFFV 6


RESULT 12
S57575
T cell receptor V-J junctional alpha chain region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57575
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified
A;Reference number: S57494
A;Accession: S57575
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-11 <BUR>
A;Cross-references: EMBL:Z49953; NID:g887510; PIDN:CAA90224.1; PID:g887511
C;Keywords: T-cell receptor

  Query Match             26.0%;  Score 20;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGRRLF 6
              ||: :|
Db          5 GGKLIF 10


RESULT 13
S26559
T-cell receptor beta chain (clone Cw3/Cas15) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C;Accession: S26559
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wi
J. Exp. Med. 176, 439-447, 1992
A;Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell recepto
A;Reference number: S26512; MUID:92364546; PMID:1380061
A;Accession: S26559
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Cross-references: EMBL:X68009
A;Experimental source: cytolytic T-lymphocyte, clone Cw3/Cas15
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor
```

```
  Query Match             35.1%;  Score 27;  DB 2;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRRLFF 7
              | ||||
Db          6 GERLFF 11


RESULT 3
S26554
T-cell receptor beta chain (clone Cw3/Cas7) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C;Accession: S26554
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wid
J. Exp. Med. 176, 439-447, 1992
A;Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor
A;Reference number: S26512; MUID:92364546; PMID:1380061
A;Accession: S26554
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Cross-references: EMBL:X68004
A;Experimental source: cytolytic T-lymphocyte, clone Cw3/Cas7
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

  Query Match             35.1%;  Score 27;  DB 2;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRRLFF 7
              | ||||
Db          6 GERLFF 11


RESULT 4
S26555
T-cell receptor beta chain (clone Cw3/1B4) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C;Accession: S26555
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wid
J. Exp. Med. 176, 439-447, 1992
A;Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor
A;Reference number: S26512; MUID:92364546; PMID:1380061
A;Accession: S26555
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Cross-references: EMBL:X68005
A;Experimental source: cytolytic T-lymphocyte, clone Cw3/1B4
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

  Query Match             35.1%;  Score 27;  DB 2;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 GRRLFF 7
              | ||||
Db          6 GERLFF 11


RESULT 5
PH0804
T-cell receptor alpha chain (L4) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0804
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0804
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60913
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             33.8%;  Score 26;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGRRLFF 7
              ||| | |
Db          7 GGRALIF 13


RESULT 6
PH0762
T-cell receptor beta chain (K1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C;Accession: PH0762
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0762
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: EMBL:X60856; NID:g52768; PIDN:CAA43246.1; PID:g52769
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             31.2%;  Score 24;  DB 2;  Length 14;
  Best Local Similarity   57.1%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 GGRRLFF 7
              ||  :||
Db          7 GGTEVFF 13


RESULT 7
PT0245
Ig heavy chain CRD3 region (clone 2-103C) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0245
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0245
A;Molecule type: DNA
A;Residues: 1-10 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

  Query Match             28.6%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 9.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGRR 4
              ||||
Db          2 GGRR 5


RESULT 8
JH0328
probursin tetradecapeptide - bovine
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-15
Perfect score: 77
Sequence: 1 GGRRLFFVKAHVRES 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	37.7	12	1	LFECPE	pyrE leader peptid
2	27	35.1	12	2	S26556	T-cell receptor be
3	27	35.1	12	2	S26554	T-cell receptor be
4	27	35.1	12	2	S26555	T-cell receptor be
5	26	33.8	14	2	PH0804	T-cell receptor al
6	24	31.2	14	2	PH0762	T-cell receptor be
7	22	28.6	10	2	PT0245	Ig heavy chain CRD
8	22	28.6	14	2	JH0328	probursin tetradec
9	22	28.6	15	2	S57577	T cell receptor V-
10	21	27.3	15	2	T03000	GTP-binding protei
11	20	26.0	10	2	A27617	triose-phosphate i
12	20	26.0	11	2	S57575	T cell receptor V-
13	20	26.0	12	2	S26559	T-cell receptor be
14	20	26.0	12	2	PH1466	T-cell receptor be
15	20	26.0	13	2	S22761	Ig lambda-2 chain
16	20	26.0	15	2	S47387	T-cell antigen rec
17	19	24.7	10	2	PH0807	T-cell receptor al
18	19	24.7	10	2	PN0165	triose-phosphate i
19	19	24.7	11	2	I65231	CCK-B gastrin rece
20	19	24.7	12	2	A33900	hydrin 1 - African
21	19	24.7	14	2	PH1332	Ig heavy chain DJ
22	19	24.7	14	2	PH1617	Ig H chain V-D-J r
23	19	24.7	15	2	S55312	TSH protein beta c
24	18	23.4	10	2	B56899	serum heterodimer,
25	18	23.4	11	2	PT0218	T-cell receptor be
26	18	23.4	11	2	PD0441	translation elonga
27	18	23.4	11	2	S45698	gamma-MSH-like pro
28	18	23.4	12	2	S26552	T-cell receptor be
29	18	23.4	12	2	S26549	T-cell receptor be
30	18	23.4	12	2	S26553	T-cell receptor be
31	18	23.4	12	2	PH1467	T-cell receptor be
32	18	23.4	12	2	PH1463	T-cell receptor be
33	18	23.4	12	2	PH1464	T-cell receptor be
34	18	23.4	12	2	PH1469	T-cell receptor be
35	18	23.4	12	2	PH1468	T-cell receptor be
36	18	23.4	13	2	JQ2309	hypothetical 1.6K
37	18	23.4	13	2	JQ2319	hypothetical 1.6K
38	18	23.4	13	2	JH0460	corticostatic pept
39	18	23.4	14	2	PH0753	T-cell receptor be
40	18	23.4	15	2	PH0808	T-cell receptor al
41	17	22.1	7	2	A28709	phosphonoacetaldeh
42	17	22.1	7	2	PT0520	T-cell receptor be
43	17	22.1	7	2	PT0667	T-cell receptor be
44	17	22.1	9	2	D24180	fibrinogen beta ch
45	17	22.1	9	2	F28854	fibrinopeptide B -

ALIGNMENTS

RESULT 1
LFECPE
pyrE leader peptide - Escherichia coli
C;Species: Escherichia coli
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: A30400; A05110; Q00495
R;Poulsen, P.; Bonekamp, F.; Jensen, K.F.
EMBO J. 3, 1783-1790, 1984
A;Title: Structure of the Escherichia coli pyrE operon and control of pyrE expression by
A;Reference number: A30400; MUID:85003588; PMID:6207018
A;Accession: A30400
A;Molecule type: DNA
A;Residues: 1-12 <POU1>
A;Cross-references: UNIPROT:P17776
R;Poulsen, P.; Jensen, K.F.; Valentin-Hansen, P.; Carlsson, P.; Lundberg, L.G.
Eur. J. Biochem. 135, 223-229, 1983
A;Title: Nucleotide sequence of the Escherichia coli pyrE gene and of the DNA in front o
A;Reference number: A05110; MUID:83287414; PMID:6349999
A;Accession: A05110
A;Molecule type: DNA
A;Residues: 1-12 <POU2>
C;Genetics:
A;Gene: pyrE-LP
A;Map position: 82 min
C;Superfamily: pyrE leader peptide

Query Match 37.7%; Score 29; DB 1; Length 12;
Best Local Similarity 70.0%; Pred. No. 54;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RLFFVKAHVR 13
||||| ||
Db 3 RLFFVCRKVR 12

RESULT 2
S26556
T-cell receptor beta chain (clone Cw3/2C3) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C;Accession: S26556
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wi
J. Exp. Med. 176, 439-447, 1992
A;Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor
A;Reference number: S26512; MUID:92364546; PMID:1380061
A;Accession: S26556
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Cross-references: EMBL:X68006
A;Experimental source: cytolytic T-lymphocyte, clone Cw3/2C3
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

XX
PS Example 1; Page 7; 19pp; English.
XX
CC The present invention relates to a method for identifying compounds for
CC treating disorders mediated by deregulation of the Mitogen Activated
CC Protein Kinase (MAPK) pathway, related to structural and/or functional
CC alterations of the endosomal/lysosomal system. The compounds are tested
CC for their ability to inhibit, on late endosomes/lysosomes, the
CC interaction of p14 with MP1 or the interaction of p14/MP1 complex with
CC components of MAPK pathway. The method is useful for identifying
CC compounds for treating disorders such as lysosomal storage/transport
CC diseases, e.g., cystic fibrosis, Nieman-Pick-disease and diseases that
CC affect lysosomal secretion, e.g., asthma or Chediak-Higashi syndrome. The
CC present sequence is a peptide fragment of murine p14 protein, which was
CC obtained in an example from the invention
XX
SQ Sequence 10 AA;

Query Match 36.8%; Score 28; DB 5; Length 10;
Best Local Similarity 83.3%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 9 KTVGFG 14
:|||||
Db 2 ETVGFG 7

RESULT 15
AAR49625
ID AAR49625 standard; peptide; 15 AA.
XX
AC AAR49625;
XX
DT 25-MAR-2003 (revised)
DT 22-AUG-1994 (first entry)
XX
DE Camel immunoglobulin variable heavy chain CDR3 peptide.
XX
KW Immunoglobulin; heavy chain; variable region; Camelid.
XX
OS Camelus dromedarius.
XX
PN WO9404678-A1.
XX
PD 03-MAR-1994.
XX
PF 18-AUG-1993; 93WO-EP002214.
XX
PR 21-AUG-1992; 92EP-00402326.
PR 21-MAY-1993; 93EP-00401310.
XX
PA (CAST/) CASTERMAN C.
PA (HAME/) HAMERS R.
XX
PI Casterman C, Hamers R;
XX
DR WPI; 1994-083195/10.
XX
PT Immunoglobulins devoid of light chains - also processes for their
PT preparation, and protein and nucleotide sequence encoding them.
XX
PS Claim 10; Page 60; 87pp; English.
XX
CC A novel immunoglobulin (Ig) is claimed which comprises two heavy (H)
CC polypeptide chains sufficient for the formation of a complete antigen
CC binding site or several such chains. The Ig is devoid of light (L)
CC polypeptide chains. The Ig may be obtd. from prokaryotic cells, esp. E.
CC coli, by: cloning a DNA or cDNA sequence coding for the VH domain of an
CC Ig devoid of L chains obtainable from e.g. lymphocytes of Camelids;
CC recovering the cloned fragment after amplification using a 5' primer
CC contg. an Xho site and a 3' primer contg. the Spe site having the
CC sequence in AAQ44383; cloning the recovered fragment is a vector;
CC transforming host cells; and recovering the expression product of the VHH
CC coding sequence. A claimed Ig comprises 4 frameworks in its variable (V)
CC region selected from:sequences in AAR49611-16 and AAR49720 for the
CC framework 1 domain; AAR49617-21 for the framework 4 domain; and/or
CC AAR49622-39 for the CDR3 domain; and/or that its constant region
CC comprises CH2 and CH3 domains comprising AA sequences selected from, for
CC the CH2 domain AAR49640-3,and for the CH3 domain AAR49444-48;and/or that
CC its hinge region comprises 0-50 AAs, esp. a sequence selected from
CC AAR49649 or AAR49650. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 15 AA;

Query Match 36.8%; Score 28; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 5.9e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 6 FLFKTVGFGG 15
| : | | ||
Db 1 FCYSTAGDGG 10

Search completed: February 22, 2005, 09:24:55
Job time : 66.6667 secs

PS Disclosure; Page 30; 588pp; English.
XX
CC The present invention describes a single chain polypeptide comprising
CC clostridial neurotoxin light and heavy chains. The single chain
CC polypeptide comprises 2 domains: the first domain is a clostridial
CC neurotoxin light chain, or its fragment or variant, which is capable of
CC cleaving one or more vesicle or plasma membrane associated proteins
CC essential to exocytosis; the second domain is a clostridial neurotoxin
CC heavy chain H-N portion, or its fragment or variant, which is capable of
CC translocating the polypeptide into a cell and/or increasing the
CC solubility of the polypeptide compared to the solubility of the first
CC domain on its own. The second domain lacks a functional C-terminal part
CC of a clostridial neurotoxin heavy chain, designated H-C, which renders
CC the polypeptide incapable of binding to cell surface receptors that are
CC the natural cell surface receptors to which native clostridial neurotoxin
CC binds. Also described is a nucleic acid molecule encoding the single
CC chain polypeptide described above. The single chain polypeptide has
CC antibacterial activity, and can be used in vaccines. The single chain
CC polypeptides can be used as positive controls for toxin assays, as
CC reagent components for the synthesis of therapeutic molecules, or for
CC developing vaccines against clostridial toxin. The polypeptides are also
CC useful as non-toxic standards for the assessment and development of in
CC vitro assays for detection of functional botulinum or tetanus neurotoxins
CC in foodstuffs or environmental samples. The present sequence is used in
CC the exemplification of the present invention.
XX
SQ Sequence 11 AA;

Query Match 38.2%; Score 29; DB 8; Length 11;
Best Local Similarity 62.5%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LFKTVGFG 14
|:| |||
Db 3 LYKKAGFG 10

RESULT 13
ADL90463
ID ADL90463 standard; peptide; 15 AA.
XX
AC ADL90463;
XX
DT 17-JUN-2004 (first entry)
XX
DE N-terminal extension peptide.
XX
KW single chain polypeptide; clostridial neurotoxin light chain;
KW clostridial neurotoxin heavy chain; Clostridium; neurotoxin; exocytosis;
KW antibacterial; vaccine; toxin assay; clostridial toxin; detection;
KW botulinum; tetanus.
XX
OS Synthetic.
XX
PN WO2004024909-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-GB003824.
XX
PR 12-SEP-2002; 2002US-00241596.
XX
PA (HEAL-) HEALTH PROTECTION AGENCY.
XX
PI Shone CC, Foster KA, Chaddock J, Marks P, Sutton MJ, Stancombe P;
PI Wayne J;
XX
DR WPI; 2004-270039/25.
XX
PT New single chain polypeptides comprising clostridial neurotoxin light and
PT heavy chains, useful as positive controls for toxin assays, or for
PT developing vaccines against clostridial toxin.
XX

PS Disclosure; Page 30; 588pp; English.
XX
CC The present invention describes a single chain polypeptide comprising
CC clostridial neurotoxin light and heavy chains. The single chain
CC polypeptide comprises 2 domains: the first domain is a clostridial
CC neurotoxin light chain, or its fragment or variant, which is capable of
CC cleaving one or more vesicle or plasma membrane associated proteins
CC essential to exocytosis; the second domain is a clostridial neurotoxin
CC heavy chain H-N portion, or its fragment or variant, which is capable of
CC translocating the polypeptide into a cell and/or increasing the
CC solubility of the polypeptide compared to the solubility of the first
CC domain on its own. The second domain lacks a functional C-terminal part
CC of a clostridial neurotoxin heavy chain, designated H-C, which renders
CC the polypeptide incapable of binding to cell surface receptors that are
CC the natural cell surface receptors to which native clostridial neurotoxin
CC binds. Also described is a nucleic acid molecule encoding the single
CC chain polypeptide described above. The single chain polypeptide has
CC antibacterial activity, and can be used in vaccines. The single chain
CC polypeptides can be used as positive controls for toxin assays, as
CC reagent components for the synthesis of therapeutic molecules, or for
CC developing vaccines against clostridial toxin. The polypeptides are also
CC useful as non-toxic standards for the assessment and development of in
CC vitro assays for detection of functional botulinum or tetanus neurotoxins
CC in foodstuffs or environmental samples. The present sequence is used in
CC the exemplification of the present invention.
XX
SQ Sequence 15 AA;

Query Match 38.2%; Score 29; DB 8; Length 15;
Best Local Similarity 62.5%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 LFKTVGFG 14
|:| |||
Db 4 LYKKAGFG 11

RESULT 14
ABB98511
ID ABB98511 standard; peptide; 10 AA.
XX
AC ABB98511;
XX
DT 25-NOV-2002 (first entry)
XX
DE Murine p14 peptide fragment #1.
XX
KW Murine; antiasthmatic; immunomodulator; asthma;
KW Mitogen Activated Protein Kinase; MAPK; endosome; lysosome; p14;
KW lysosomal storage disease; lysosomal transport disease; cystic fibrosis;
KW Nieman-Pick-disease; lysosomal secretion; Chediak-Higashi syndrome.
XX
OS Mus musculus.
XX
PN EP1233271-A1.
XX
PD 21-AUG-2002.
XX
PF 16-FEB-2001; 2001EP-00103792.
XX
PR 16-FEB-2001; 2001EP-00103792.
XX
PA (HUBE/) HUBER L.
PA (WUND/) WUNDERLICH W.
PA (FIAL/) FIALKA I.
XX
PI Huber L, Wunderlich W, Fialka I;
XX
DR WPI; 2002-668424/72.
XX
PT Identifying compounds for treating disorders mediated by deregulation of
PT mitogen activated protein kinase pathway related to structural and/or
PT functional alterations of endosomal/lysosomal system.

XX
PT Analyzing cleavage of polymer, by providing polymer sample, incubating
PT the sample with labeled isotope for cleavage at potential cleavage site,
PT and analyzing the masses of any uncleaved fragments by mass spectrometry.
XX
PS Example 8; Page 33; 73pp; English.
XX
CC The present invention describes a method (M1) for analysing cleavage of a
CC polymer. M1 comprises: (a) providing a sample of the polymer, a portion
CC of the polymer molecules having been labeled at a position on one side of
CC the potential cleavage site with a first isotopic label and a portion of
CC the polymer molecules having been labeled at a position on the opposite
CC side of the potential cleavage site with a second isotopic label; (b)
CC incubating the sample under conditions suitable for cleavage at the
CC potential cleavage site; and (c) analysing the mass(es) of any cleaved
CC fragments by mass spectrometry and thereby determining whether and/or
CC where cleavage has taken place. M1 is useful for analysing cleavage of a
CC polymer, where the polymer is a linear polymer, and comprises a peptide
CC or protein. Methods from the present invention can be used in discovering
CC new or improved synthetic substrates for both known and unknown enzymes,
CC e.g. enzymes identified from the human genome. The methods are also
CC useful to identify the sequence origin, and in screening methods to
CC identify new substrates for enzymes, in positional peptide scanning
CC libraries, in in vivo/ex vivo/in vitro peptide, and in assaying methods
CC for oligonucleotide or peptide sequencing and in measuring differential
CC protein expression. The methods are useful for monitoring the cleavage of
CC polypeptides or polynucleotides, and for determining optimal polymer
CC substrates. ABP57505 to ABP57605 represent peptides used in the
CC exemplification of the present invention
XX
SQ Sequence 10 AA;

Query Match 38.2%; Score 29; DB 6; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAFLFK 9
||||||
Db 1 AAFLFK 6

RESULT 11
ABP57592
ID ABP57592 standard; peptide; 11 AA.
XX
AC ABP57592;
XX
DT 28-APR-2003 (first entry)
XX
DE Differentially isotopically labelled (DiMaS) peptide #78.
XX
KW Mass spectrometry; polymer; analysis; cleavage; substrate specificity;
KW isotope; protease.
XX
OS Synthetic.
XX
PN WO2003001206-A1.
XX
PD 03-JAN-2003.
XX
PF 25-JUN-2002; 2002WO-GB002921.
XX
PR 26-JUN-2001; 2001GB-00015581.
XX
PA (GLAX) GLAXO GROUP LTD.
XX
PI Mckeown SC;
XX
DR WPI; 2003-184066/18.
XX
PT Analyzing cleavage of polymer, by providing polymer sample, incubating
PT the sample with labeled isotope for cleavage at potential cleavage site,
PT and analyzing the masses of any uncleaved fragments by mass spectrometry.
XX
PS Example 8; Page 36; 73pp; English.
XX
CC The present invention describes a method (M1) for analysing cleavage of a
CC polymer. M1 comprises: (a) providing a sample of the polymer, a portion
CC of the polymer molecules having been labeled at a position on one side of
CC the potential cleavage site with a first isotopic label and a portion of
CC the polymer molecules having been labeled at a position on the opposite
CC side of the potential cleavage site with a second isotopic label; (b)
CC incubating the sample under conditions suitable for cleavage at the
CC potential cleavage site; and (c) analysing the mass(es) of any cleaved
CC fragments by mass spectrometry and thereby determining whether and/or
CC where cleavage has taken place. M1 is useful for analysing cleavage of a
CC polymer, where the polymer is a linear polymer, and comprises a peptide
CC or protein. Methods from the present invention can be used in discovering
CC new or improved synthetic substrates for both known and unknown enzymes,
CC e.g. enzymes identified from the human genome. The methods are also
CC useful to identify the sequence origin, and in screening methods to
CC identify new substrates for enzymes, in positional peptide scanning
CC libraries, in in vivo/ex vivo/in vitro peptide, and in assaying methods
CC for oligonucleotide or peptide sequencing and in measuring differential
CC protein expression. The methods are useful for monitoring the cleavage of
CC polypeptides or polynucleotides, and for determining optimal polymer
CC substrates. ABP57505 to ABP57605 represent peptides used in the
CC exemplification of the present invention
XX
SQ Sequence 11 AA;

Query Match 38.2%; Score 29; DB 6; Length 11;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAFLFK 9
||||||
Db 2 AAFLFK 7

RESULT 12
ADL90462
ID ADL90462 standard; peptide; 11 AA.
XX
AC ADL90462;
XX
DT 17-JUN-2004 (first entry)
XX
DE N-terminal extension peptide.
XX
KW single chain polypeptide; clostridial neurotoxin light chain;
KW clostridial neurotoxin heavy chain; Clostridium; neurotoxin; exocytosis;
KW antibacterial; vaccine; toxin assay; clostridial toxin; detection;
KW botulinum; tetanus.
XX
OS Synthetic.
XX
PN WO2004024909-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-GB003824.
XX
PR 12-SEP-2002; 2002US-00241596.
XX
PA (HEAL-) HEALTH PROTECTION AGENCY.
XX
PI Shone CC, Foster KA, Chaddock J, Marks P, Sutton MJ, Stancombe P;
PI Wayne J;
XX
DR WPI; 2004-270039/25.
XX
PT New single chain polypeptides comprising clostridial neurotoxin light and
PT heavy chains, useful as positive controls for toxin assays, or for
PT developing vaccines against clostridial toxin.
XX

```
CC   polynucleotides and specific antibodies. Immunotherapy; Anti-HIV. The
CC   compounds are useful for inducing an immune response and administering
CC   immunotherapy to a subject with AIDS. The present sequence represents a
CC   specific example of a HIV SL9 peptide epitope derived compound.
XX
SQ   Sequence 9 AA;

  Query Match             39.5%;  Score 30;  DB 8;  Length 9;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 PLFKTVG 12
              ||::|||
Db          1 PLYRTVG 7


RESULT 8
ABR44302
ID   ABR44302 standard; peptide; 15 AA.
XX
AC   ABR44302;
XX
DT   27-JUN-2003  (first entry)
XX
DE   Human L-asparaginase 24.53 N-terminal peptide sequence #SEQ ID 7.
XX
KW   Human; L-asparaginase; 24.53; diabetes; tumour.
XX
OS   Homo sapiens.
XX
PN   CN1360049-A.
XX
PD   24-JUL-2002.
XX
PF   20-DEC-2000; 2000CN-00135126.
XX
PR   20-DEC-2000; 2000CN-00135126.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-733674/80.
XX
PT   A human L-asparaginase 24.53 polypeptide L-asparaginase 24.53 and the
PT   polynucleotide encoding it.
XX
PS   Example 6; Page 18 (disclosure); 32pp; Chinese.
XX
CC   The invention relates to a human L-asparaginase polypeptide designated
CC   24.53, and the polynucleotide encoding it. Also disclosed is the method
CC   for preparing the polypeptide using DNA recombination techniques. The
CC   application of the polypeptide is in treating diseases including diabetes
CC   and tumours. The current sequence represents the human L-asparaginase N-
CC   terminal peptide sequence
XX
SQ   Sequence 15 AA;

  Query Match             39.5%;  Score 30;  DB 5;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 2.6e+02;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    4;  Gaps    1;

Qy          1 MARAAFLFKTVGFG 14
              | |||    |||:|
Db          1 MVRAA----TVGYG 10


RESULT 9
ABP55159
ID   ABP55159 standard; peptide; 15 AA.
XX
AC   ABP55159;
XX
DT   05-FEB-2003  (first entry)
XX
DE   Mouse myoglobulin I-111-18.48 N-terminal peptide.
XX
KW   Mouse; myoglobulin I-111-18.48; diabetes; tumour; antidiabetic;
KW   antitumour.
XX
OS   Mus sp.
XX
PN   CN1343726-A.
XX
PD   10-APR-2002.
XX
PF   19-SEP-2000; 2000CN-00125213.
XX
PR   19-SEP-2000; 2000CN-00125213.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-548894/59.
XX
PT   A novel mouse myoglobulin I-111-18.48 polypeptide, useful for treating
PT   several diseases e.g. diabetes and tumors.
XX
PS   Example 6; Page 19 (Disclosure); 31pp; Chinese.
XX
CC   The present invention relates to novel mouse myoglobulin I-111-18.48 (see
CC   ABP55158) and to a polynucleotide (see ABV75939) encoding it. The protein
CC   is useful for the treatment of several diseases, e.g. diabetes and
CC   tumours. The present sequence is an N-terminal peptide fragment of the
CC   myoglobulin, which was used in an example from the invention
XX
SQ   Sequence 15 AA;

  Query Match             39.5%;  Score 30;  DB 5;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 2.6e+02;
  Matches    6;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 MARAAFLFKTVGFG 14
              |   | |::||   |
Db          1 MVEKALLYRTVATG 14


RESULT 10
ABP57556
ID   ABP57556 standard; peptide; 10 AA.
XX
AC   ABP57556;
XX
DT   28-APR-2003  (first entry)
XX
DE   Differentially isotopically labelled (DiMaS) peptide #42.
XX
KW   Mass spectrometry; polymer; analysis; cleavage; substrate specificity;
KW   isotope; protease.
XX
OS   Synthetic.
XX
PN   WO2003001206-A1.
XX
PD   03-JAN-2003.
XX
PF   25-JUN-2002; 2002WO-GB002921.
XX
PR   26-JUN-2001; 2001GB-00015581.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Mckeown SC;
XX
DR   WPI; 2003-184066/18.
```

PR 15-JUN-1999; 99WO-IL000329.
XX
PA (HORN/) HORNIK V.
PA (AFAR/) AFARGAN M M.
PA (GELL/) GELLERMAN G.
XX
PI Hornik V, Afargan MM, Gellerman G;
XX
DR WPI; 2002-681319/73.
XX
PT New backbone cyclized somatostatin analogs are e.g. useful in the
PT treatment of atherosclerosis, autoimmune diseases and cancers.
XX
PS Disclosure; Page 12; 30pp; English.
XX
CC The present invention describes backbone cyclised somatostatin analogues
CC (I) that incorporates at least one building unit containing one nitrogen
CC atom of the peptide backbone connected to a bridging group (comprising an
CC amide, thioether, thioester or disulfide) where at least one building
CC unit is connected via the bridging group to form a cyclic structure with
CC a moiety selected from the group consisting of a second building unit,
CC the side chain of an amino acid residue of the sequence or the N-terminal
CC amino acid residue. (I) has antiarteriosclerotic, immunosuppressive,
CC cytostatic, antidiabetic, antiinflammatory and analgesic activities, and
CC can be used as a somatostatin receptor ligand. (I) are useful in the
CC treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-
CC associated complications, endocrine disorders, inflammation,
CC gastrointestinal disorders, pancreatitis, post-surgical pain, and
CC restenosis. (I) can also be used in the diagnosis of cancer, by imaging
CC the existence of metastases, it being labeled with a detectable probe.
CC The present sequence represents a multiple parallel synthesis
CC somatostatin analogue from the present invention
XX
SQ Sequence 9 AA;

Query Match 42.1%; Score 32; DB 5; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.8e+06;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 6 PLPKTVGP 13
| :|||||
Db 2 PXWKTVGP 9

RESULT 6
AAW42262
ID AAW42262 standard; peptide; 15 AA.
XX
AC AAW42262;
XX
DT 08-APR-1998 (first entry)
XX
DE Biotinylated cross-linked interleukin-8 15-mer peptide ligand 2.
XX
KW Bacteriophage peptide library; peptide epitope; therapeutic target;
KW variegated compound library; interleukin-8; IL-8.
XX
OS Synthetic.
XX
PN WO9735194-A2.
XX
PD 25-SEP-1997.
XX
PF 21-MAR-1997; 97WO-US004176.
XX
PR 21-MAR-1996; 96US-00622338.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Forster AC;
XX
DR WPI; 1997-480355/44.
XX
PT Identifying compounds which interact with target molecules - using
PT enantiomers of the target molecules and testing of enantiomers of
PT selected compounds.
XX
PS Disclosure; Fig 6; 89pp; English.
XX
CC 15-mer peptides AAW42261-77 are identified as ligands of a biotinylated,
CC cross-linked interleukin-8 (IL-8) target, using the method of the
CC invention. This novel method identifies compounds which interact with a
CC target molecule, and comprises contacting a screening molecule with a
CC variegated compound library, where the screening molecule comprises solid
CC target molecule, or the enantiomer if the target molecule is chiral.
CC Compounds which have a desired interaction with the target molecule are
CC selected, and the ability of their enantiomer to interact with the target
CC molecule is tested. Ligands for a target protein can be identified by
CC combining a D-enantiomer of a target protein (a D-target protein), and a
CC variegated compound library, and then selecting one or more compounds
CC from the library which have a desired binding interaction with the D-
CC target protein. The methods can be used for identifying agonists or
CC antagonists of targets such as receptors, enzymes, DNA binding proteins
CC or signal transduction proteins. The methods can provide a structurally
CC selective approach in addition to scoring for interaction of functional
CC groups. They provide a powerful selection method that allows for the
CC production of ligands with the same diversity as peptides but with the
CC greatly improved pharmacokinetic profiles needed for drug activity
XX
SQ Sequence 15 AA;

Query Match 40.8%; Score 31; DB 2; Length 15;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 RAAFLFKTVG 12
|:|| | :||
Db 5 RSAFRFSSVG 14

RESULT 7
ADO44298
ID ADO44298 standard; peptide; 9 AA.
XX
AC ADO44298;
XX
DT 29-JUL-2004 (first entry)
XX
DE HIV SL9 epitope derived compound 4.
XX
KW SL9; gag; HIV; immunotherapy; anti-HIV; AIDS; epitope.
XX
OS Human immunodeficiency virus 1.
XX
PN WO2004039324-A2.
XX
PD 13-MAY-2004.
XX
PF 24-OCT-2003; 2003WO-US033977.
XX
PR 28-OCT-2002; 2002US-0422038P.
XX
PA (GENZ) GENZYME CORP.
PA (MASS-) MASSACHUSETTS GEN HOSPITAL.
XX
PI Nicolette CA, Walker BD;
XX
DR WPI; 2004-376049/35.
DR N-PSDB; ADO44299.
XX
PT New anti-HIV (SL9) compounds, useful for inducing an immune response and
PT administering immunotherapy to a subject with AIDS.
XX
PS Claim 1; SEQ ID NO 7; 58pp; English.
XX
CC The invention relates to anti-HIV (SL9) compounds and encoding

```
CC   amide, thioether, thioester or disulfide) where at least one building
CC   unit is connected via the bridging group to form a cyclic structure with
CC   a moiety selected from the group consisting of a second building unit,
CC   the side chain of an amino acid residue of the sequence or the N-terminal
CC   amino acid residue. (I) has antiarteriosclerotic, immunosuppressive,
CC   cytostatic, antidiabetic, antiinflammatory and analgesic activities, and
CC   can be used as a somatostatin receptor ligand. (I) are useful in the
CC   treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-
CC   associated complications, endocrine disorders, inflammation,
CC   gastrointestinal disorders, pancreatitis, post-surgical pain, and
CC   restenosis. (I) can also be used in the diagnosis of cancer, by imaging
CC   the existence of metastases, it being labeled with a detectable probe.
CC   The present sequence represents a multiple parallel synthesis
CC   somatostatin analogue from the present invention
XX
SQ   Sequence 9 AA;

  Query Match             43.4%;  Score 33;  DB 5;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 FLFKTVGF 13
              | :|||||
Db          2 FFWKTVGF 9


RESULT 4
ABP53382
ID   ABP53382 standard; peptide; 9 AA.
XX
AC   ABP53382;
XX
DT   19-NOV-2002  (first entry)
XX
DE   Multiple parallel synthesis somatostatin analogue #13.
XX
KW   Backbone cyclised somatostatin analogue; somatostatin; SRIF; analgesic;
KW   somatotropin release inhibiting factor; somatostatin receptor subtype;
KW   synthesis; antiarteriosclerotic; immunosuppressive; cytostatic; cancer;
KW   antidiabetic; antiinflammatory; somatostatin receptor ligand;
KW   atherosclerosis; autoimmune disease; diabetic-associated complication;
KW   endocrine disorder; inflammation; gastrointestinal disorder; restenosis;
KW   pancreatitis; post-surgical pain.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= bAla
FT                   /note= "beta-alanine"
FT   Misc-difference 4
FT                   /note= "D form residue"
FT   Misc-difference 9
FT                   /note= "D form residue"
XX
PN   US2002052315-A1.
XX
PD   02-MAY-2002.
XX
PF   13-DEC-2000; 2000US-00734583.
XX
PR   19-JUN-1998;   98US-00100360.
PR   02-DEC-1998;   98US-00203389.
PR   15-JUN-1999;   99WO-IL000329.
XX
PA   (HORN/) HORNIK V.
PA   (AFAR/) AFARGAN M M.
PA   (GELL/) GELLERMAN G.
XX
PI   Hornik V,  Afargan MM,  Gellerman G;
XX
DR   WPI; 2002-681319/73.
XX
PT   New backbone cyclized somatostatin analogs are e.g. useful in the
PT   treatment of atherosclerosis, autoimmune diseases and cancers.
XX
PS   Disclosure; Page 12; 30pp; English.
XX
CC   The present invention describes backbone cyclised somatostatin analogues
CC   (I) that incorporates at least one building unit containing one nitrogen
CC   atom of the peptide backbone connected to a bridging group (comprising an
CC   amide, thioether, thioester or disulfide) where at least one building
CC   unit is connected via the bridging group to form a cyclic structure with
CC   a moiety selected from the group consisting of a second building unit,
CC   the side chain of an amino acid residue of the sequence or the N-terminal
CC   amino acid residue. (I) has antiarteriosclerotic, immunosuppressive,
CC   cytostatic, antidiabetic, antiinflammatory and analgesic activities, and
CC   can be used as a somatostatin receptor ligand. (I) are useful in the
CC   treatment of atherosclerosis, autoimmune diseases, cancers, diabetic-
CC   associated complications, endocrine disorders, inflammation,
CC   gastrointestinal disorders, pancreatitis, post-surgical pain, and
CC   restenosis. (I) can also be used in the diagnosis of cancer, by imaging
CC   the existence of metastases, it being labeled with a detectable probe.
CC   The present sequence represents a multiple parallel synthesis
CC   somatostatin analogue from the present invention
XX
SQ   Sequence 9 AA;

  Query Match             42.1%;  Score 32;  DB 5;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+06;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 FLFKTVGF 13
              | :|||||
Db          2 FYWKTVGF 9


RESULT 5
ABP53384
ID   ABP53384 standard; peptide; 9 AA.
XX
AC   ABP53384;
XX
DT   19-NOV-2002  (first entry)
XX
DE   Multiple parallel synthesis somatostatin analogue #15.
XX
KW   Backbone cyclised somatostatin analogue; somatostatin; SRIF; analgesic;
KW   somatotropin release inhibiting factor; somatostatin receptor subtype;
KW   synthesis; antiarteriosclerotic; immunosuppressive; cytostatic; cancer;
KW   antidiabetic; antiinflammatory; somatostatin receptor ligand;
KW   atherosclerosis; autoimmune disease; diabetic-associated complication;
KW   endocrine disorder; inflammation; gastrointestinal disorder; restenosis;
KW   pancreatitis; post-surgical pain.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= bAla
FT                   /note= "beta-alanine"
FT   Modified-site   3
FT                   /note= "cyclohexyl glycine"
FT   Misc-difference 4
FT                   /note= "D form residue"
FT   Misc-difference 9
FT                   /note= "D form residue"
XX
PN   US2002052315-A1.
XX
PD   02-MAY-2002.
XX
PF   13-DEC-2000; 2000US-00734583.
XX
PR   19-JUN-1998;   98US-00100360.
PR   02-DEC-1998;   98US-00203389.
```

CC the autoimmune disease. This peptide is an internal peptide of reovirus
CC type 3 sigma 2 protein and is implicated as a foreign epitope involved in
CC the aetiology or in remissions of multiple sclerosis. It has been shown
CC capable of inducing the proliferation of autoreactive T-cell clones
CC isolated from multiple sclerosis patients. (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 76; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARAAFLFKTVGFGG 15
||||||||||||||||
Db 1 MARAAFLFKTVGFGG 15

RESULT 2
AAY34138
ID AAY34138 standard; peptide; 15 AA.
XX
AC AAY34138;
XX
DT 30-NOV-1999 (first entry)
XX
DE Variant human potassium channel pore domain peptide sequence 5.
XX
KW Potassium channel; ataxia; arrhythmia; epilepsy; Bartter's syndrome;
KW cardiovascular disorder; CNS disorder; renal disorder.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9943696-A1.
XX
PD 02-SEP-1999.
XX
PF 22-FEB-1999; 99WO-US003826.
XX
PR 25-FEB-1998; 98US-0076687P.
PR 07-AUG-1998; 98US-0095836P.
PR 19-JAN-1999; 99US-0116448P.
XX
PA (AXYS-) AXYS PHARM INC.
XX
PI Miller AP, Curran ME, Hu P, Rutter M, Wang J;
XX
DR WPI; 1999-527591/44.
XX
PT New nucleic acids encoding mammalian K+Hnov potassium channel proteins,
PT useful for the diagnosis and treatment of episodic ataxia with myokymia,
PT cardiac arrhythmia, epilepsy and Bartter's syndrome.
XX
PS Example 1; Page 32; 112pp; English.
XX
CC This sequence represents a variant human potassium channel pore domain
CC peptide sequence used in the identification and isolation of human K+Hnov
CC cDNAs (AAZ11897-Z11915). K+Hnov proteins have a high degree of homology
CC to known potassium channels and may be alpha subunits, which form the
CC functional channel, or accessory subunits that act to modulate the
CC channel activity. K+Hnov cDNAs were isolated by extension of expressed
CC sequence tags (ESTs) which were related but not identical to known human
CC potassium channels. Potential polymorphisms detected as sequence variants
CC between multiple independent clones. Potassium channels have critical
CC roles in various cell types and biochemical pathways. Defective potassium
CC channels are known to cause four human diseases: episodic ataxia with
CC myokymia; cardiac arrhythmia (long QT syndrome); epilepsy; and Bartter's
CC syndrome. As potassium channels are critical components of virtually all
CC cells, it is likely that abnormal potassium channels are also implicated
CC in certain renal, cardiovascular and central nervous system (CNS)
CC disorders. Nucleotides encoding K+Hnov proteins may be used for
CC identifying homologous or related proteins and the DNA sequences encoding
CC them. They may be used to produce compositions that modulate the
CC expression and function of the K+Hnov protein and in studying the
CC biochemical pathways associated with it. They may also be used for the
CC recombinant production of K+Hnov protein in fermentation cultures.
CC Additionally, such nucleotides may be used in gene therapy protocols for
CC the treatment of diseases associated with abnormal potassium channels
XX
SQ Sequence 15 AA;

Query Match 46.7%; Score 35.5; DB 2; Length 15;
Best Local Similarity 53.3%; Pred. No. 26;
Matches 8; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 6 FLFK-----TVGFGG 15
||| |:||||
Db 1 FLFSIEVQVTIGFGG 15

RESULT 3
ABP53383
ID ABP53383 standard; peptide; 9 AA.
XX
AC ABP53383;
XX
DT 19-NOV-2002 (first entry)
XX
DE Multiple parallel synthesis somatostatin analogue #14.
XX
KW Backbone cyclised somatostatin analogue; somatostatin; SRIF; analgesic;
KW somatotropin release inhibiting factor; somatostatin receptor subtype;
KW synthesis; antiarteriosclerotic; immunosuppressive; cytostatic; cancer;
KW antidiabetic; antiinflammatory; somatostatin receptor ligand;
KW atherosclerosis; autoimmune disease; diabetic-associated complication;
KW endocrine disorder; inflammation; gastrointestinal disorder; restenosis;
KW pancreatitis; post-surgical pain.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= bAla
FT /note= "beta-alanine"
FT Misc-difference 4
FT /note= "D form residue"
FT Misc-difference 9
FT /note= "D form residue"
XX
PN US2002052315-A1.
XX
PD 02-MAY-2002.
XX
PF 13-DEC-2000; 2000US-00734583.
XX
PR 19-JUN-1998; 98US-00100360.
PR 02-DEC-1998; 98US-00203389.
PR 15-JUN-1999; 99WO-IL000329.
XX
PA (HORN/) HORNIK V.
PA (AFAR/) AFARGAN M M.
PA (GELL/) GELLERMAN G.
XX
PI Hornik V, Afargan MM, Gellerman G;
XX
DR WPI; 2002-681319/73.
XX
PT New backbone cyclized somatostatin analogs are e.g. useful in the
PT treatment of atherosclerosis, autoimmune diseases and cancers.
XX
PS Disclosure; Page 12; 30pp; English.
XX
CC The present invention describes backbone cyclised somatostatin analogues
CC (I) that incorporates at least one building unit containing one nitrogen
CC atom of the peptide backbone connected to a bridging group (comprising an

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-14
Perfect score: 76
Sequence: 1 MARAAFLPKTVGFGG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	76	100.0	15	2	AAW04854	Aaw04854 Internal
2	35.5	46.7	15	2	AAY34138	Aay34138 Variant h
3	33	43.4	9	5	ABP53383	Abp53383 Multiple
4	32	42.1	9	5	ABP53382	Abp53382 Multiple
5	32	42.1	9	5	ABP53384	Abp53384 Multiple
6	31	40.8	15	2	AAW42262	Aaw42262 Biotinyla
7	30	39.5	9	8	ADO44298	Ado44298 HIV SL9 e
8	30	39.5	15	5	ABR44302	Abr44302 Human L-a
9	30	39.5	15	5	ABP55159	Abp55159 Mouse myo
10	29	38.2	10	6	ABP57556	Abp57556 Different
11	29	38.2	11	6	ABP57592	Abp57592 Different
12	29	38.2	11	8	ADL90462	Adl90462 N-termina
13	29	38.2	15	8	ADL90463	Adl90463 N-termina
14	28	36.8	10	5	ABB98511	Abb98511 Murine p1
15	28	36.8	15	2	AAR49625	Aar49625 Camel imm
16	27	35.5	7	3	AAB19614	Aab19614 Hepatocyt
17	27	35.5	8	2	AAW56715	Aaw56715 Aspergill
18	27	35.5	9	6	ABJ41536	Abj41536 151P3D4 c
19	27	35.5	9	6	ABJ44158	Abj44158 151P3D4 c
20	27	35.5	9	6	ABJ44229	Abj44229 151P3D4 c
21	27	35.5	9	6	ABJ45425	Abj45425 151P3D4 c
22	27	35.5	9	6	ABJ49519	Abj49519 151P3D4 c
23	27	35.5	9	6	ABJ44588	Abj44588 151P3D4 c
24	27	35.5	9	6	ABJ45584	Abj45584 151P3D4 c
25	27	35.5	9	6	ABJ46858	Abj46858 151P3D4 c
26	27	35.5	9	6	ABJ39877	Abj39877 151P3D4 c
27	27	35.5	9	6	ABJ41104	Abj41104 151P3D4 c
28	27	35.5	9	6	ABJ42770	Abj42770 151P3D4 c
29	27	35.5	9	6	ABJ41943	Abj41943 151P3D4 c
30	27	35.5	9	6	ABJ43336	Abj43336 151P3D4 c
31	27	35.5	9	6	ABJ39882	Abj39882 151P3D4 c
32	27	35.5	9	6	ABJ39914	Abj39914 151P3D4 c
33	27	35.5	9	6	ABJ40765	Abj40765 151P3D4 c
34	27	35.5	9	6	ABJ43142	Abj43142 151P3D4 c
35	27	35.5	9	6	ABJ44908	Abj44908 151P3D4 c
36	27	35.5	9	6	ABJ46012	Abj46012 151P3D4 c
37	27	35.5	9	6	ABJ48740	Abj48740 151P3D4 c
38	27	35.5	9	6	ABJ44710	Abj44710 151P3D4 c
39	27	35.5	9	6	ABJ46636	Abj46636 151P3D4 c
40	27	35.5	9	6	ABJ48261	Abj48261 151P3D4 c
41	27	35.5	9	6	ABJ49726	Abj49726 151P3D4 c
42	27	35.5	9	6	ABJ45324	Abj45324 151P3D4 c
43	27	35.5	9	6	ABJ48040	Abj48040 151P3D4 c
44	27	35.5	9	6	ABJ41916	Abj41916 151P3D4 c
45	27	35.5	9	6	ABJ44665	Abj44665 151P3D4 c

ALIGNMENTS

RESULT 1
AAW04854
ID AAW04854 standard; peptide; 15 AA.
XX
AC AAW04854;
XX
DT 16-OCT-2003 (revised)
DT 18-FEB-1997 (first entry)
XX
DE Internal fragment of reovirus type 3 sigma 2 protein.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Reovirus sp.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 2; Page 46; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with

```
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Varanus gilleni (Pygmy mulga monitor).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX   NCBI_TaxID=169840;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ast J.C.;
RT   "Mitochondrial DNA evidence and evolution in Varanoidea (Squamata).";
RL   Cladistics 17:211-226(2001).
DR   EMBL; AF407499; AAL10051.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1340 MW;  CF6DEE80C733640D CRC64;

  Query Match             28.9%;  Score 22;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 7.4e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MARAAFLFKT 10
              |  | :|| |
Db          1 MTLARWLFST 10


Search completed: February 22, 2005, 09:38:05
Job time : 54.6667 secs
```

```
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Cuticle protein 30 (LM-30) (LM-ACP 30) (Fragment).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC   Acridoidea; Acrididae; Oedipodinae; Locusta.
OX   NCBI_TaxID=7004;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=86108304; PubMed=3943519;
RA   Hoejrup P., Andersen S.O., Roepstorff P.;
RT   "Isolation, characterization, and N-terminal sequence studies of
RT   cuticular proteins from the migratory locust, Locusta migratoria.";
RL   Eur. J. Biochem. 154:153-159(1986).
CC   -!- FUNCTION: Component of the cuticle of migratory locust which
CC       contains more than 100 different structural proteins.
DR   PIR; H24802; H24802.
KW   Cuticle; Direct protein sequencing; Structural protein.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  969 MW;  4973E36B58772877 CRC64;

  Query Match             28.9%;  Score 22;  DB 1;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 6.8e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         11 VGFGG 15
              :|:||
Db          5 LGYGG 9


RESULT 12
Q85J75
ID   Q85J75      PRELIMINARY;      PRT;    10 AA.
AC   Q85J75;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Varanus brevicauda (Short-tailed monitor).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Varanidae; Varanus.
OX   NCBI_TaxID=62038;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ast J.C.;
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY264940; AAP22709.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1225 MW;  5DEE80C733640DD7 CRC64;

  Query Match             28.9%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.8e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MARAAFLFKT 10
              |  | :|| |
Db          1 MTLARWLFST 10


RESULT 13
Q9TG32
ID   Q9TG32      PRELIMINARY;      PRT;    10 AA.
AC   Q9TG32;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Ophisaurus ventralis (Eastern glass lizard).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguidae;
OC   Ophisaurus.
OX   NCBI_TaxID=102195;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99343613; PubMed=10413621; DOI=10.1006/mpev.1999.0615;
RA   Macey J.R., Schulte J.A. II, Larson A., Tuniyev B.S., Orlov N.,
RA   Papenfuss T.J.;
RT   "Molecular phylogenetics, tRNA evolution, and historical biogeography
RT   in anguid lizards and related taxonomic families.";
RL   Mol. Phylogenet. Evol. 12:250-272(1999).
DR   EMBL; AF085626; AAD51568.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1225 MW;  5DEE80C733640DD7 CRC64;

  Query Match             28.9%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.8e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MARAAFLFKT 10
              |  | :|| |
Db          1 MTLARWLFST 10


RESULT 14
Q9TG47
ID   Q9TG47      PRELIMINARY;      PRT;    10 AA.
AC   Q9TG47;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Ophisaurus koellikeri.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguidae;
OC   Ophisaurus.
OX   NCBI_TaxID=102194;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99343613; PubMed=10413621; DOI=10.1006/mpev.1999.0615;
RA   Macey J.R., Schulte J.A. II, Larson A., Tuniyev B.S., Orlov N.,
RA   Papenfuss T.J.;
RT   "Molecular phylogenetics, tRNA evolution, and historical biogeography
RT   in anguid lizards and related taxonomic families.";
RL   Mol. Phylogenet. Evol. 12:250-272(1999).
DR   EMBL; AF085621; AAD51553.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1225 MW;  5DEE80C733640DD7 CRC64;

  Query Match             28.9%;  Score 22;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.8e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MARAAFLFKT 10
              |  | :|| |
Db          1 MTLARWLFST 10


RESULT 15
Q94VH7
ID   Q94VH7      PRELIMINARY;      PRT;    11 AA.
AC   Q94VH7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
```

```
ID   Q7CPD2      PRELIMINARY;      PRT;    15 AA.
AC   Q7CPD2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Spot 42 RNA, inhibition of DNA synthesis.
GN   Name=spf; OrderedLocusNames=STM4000;
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
DR   EMBL; AE008887; AAL22839.1; -.
KW   Complete proteome.
SQ   SEQUENCE   15 AA;  1725 MW;  5A51DAF170EA661E CRC64;

  Query Match             31.6%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 LFKTVGFG 14
              |   :|||
Db          8 LLHVIGFG 15


RESULT 8
Q70F01
ID   Q70F01      PRELIMINARY;      PRT;    10 AA.
AC   Q70F01;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Calpastatin type 2 (Fragment).
GN   Name=CAST;
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21531263; PubMed=11673859; DOI=10.1006/abbi.2001.2546;
RA   Parr T., Sensky P.L., Bardsley R.G., Buttery P.J.;
RT   "Calpastatin expression in porcine cardiac and skeletal muscle and
RT   partial gene structure.";
RL   Arch. Biochem. Biophys. 395:1-13(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Parr T.;
RL   Submitted (SEP-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ583410; CAE47431.1; -.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1291 MW;  CFF912436365BDD9 CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 4.5e+03;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 MARAAFLFKT 10
              || |:: :||
Db          1 MAFASWWYKT 10


RESULT 9
Q9TG68
ID   Q9TG68      PRELIMINARY;      PRT;    10 AA.
AC   Q9TG68;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Gerrhonotus liocephalus.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Anguidae;
OC   Gerrhonotus.
OX   NCBI_TaxID=76654;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99343613; PubMed=10413621; DOI=10.1006/mpev.1999.0615;
RA   Macey J.R., Schulte J.A. II, Larson A., Tuniyev B.S., Orlov N.,
RA   Papenfuss T.J.;
RT   "Molecular phylogenetics, tRNA evolution, and historical biogeography
RT   in anguid lizards and related taxonomic families.";
RL   Mol. Phylogenet. Evol. 12:250-272(1999).
DR   EMBL; AF085614; AAD51532.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1213 MW;  85EE80C733640DC1 CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 4.5e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 MARAAFLFKT 10
              |  | :|| |
Db          1 MTTARWLFST 10


RESULT 10
Q7M4Z7
ID   Q7M4Z7      PRELIMINARY;      PRT;    15 AA.
AC   Q7M4Z7;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Cytochrome c2 (Fragment).
OS   Fusarium sporotrichioides.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Hypocreomycetidae; Hypocreales; mitosporic Hypocreales; Fusarium.
OX   NCBI_TaxID=5514;
RN   [1]
RP   SEQUENCE.
RA   Chow L.P., Fukaya N., Sugiura Y., Ueno Y., Tabuchi K., Tsugita A.;
RL   Submitted (OCT-1994) to the PIR data bank.
DR   PIR; PA0087; PA0087.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1576 MW;  37EE5E40C4E886DD CRC64;

  Query Match             30.3%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 6.6e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 RAAFLFKT 10
              : | ||||
Db          7 KGANLFKT 14


RESULT 11
CU30_LOCMI
ID   CU30_LOCMI     STANDARD;      PRT;    10 AA.
AC   P11735;
DT   01-OCT-1989 (Rel. 12, Created)
```

```
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1459 MW;  03FA0D6AA88B9A30 CRC64;

  Query Match             38.8%;  Score 29.5;  DB 2;  Length 14;
  Best Local Similarity   60.0%;  Pred. No. 4e+02;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          1 MARAAFLFKTVGFGG 15
              ||||  : ||||  |
Db          1 MARARSI-KTVGLVG 14


RESULT 3
Q9S8Q8
ID   Q9S8Q8      PRELIMINARY;      PRT;    15 AA.
AC   Q9S8Q8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Phospholipase D (EC 3.1.4.4) (Fragment).
OS   Ricinus communis (Castor bean).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC   eurosids I; Malpighiales; Euphorbiaceae; Acalyphoideae; Acalypheae;
OC   Ricinus.
OX   NCBI_TaxID=3988;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=94029022; PubMed=8215453; DOI=10.1006/abbi.1993.1541;
RA   Wang X., Dyer J.H., Zheng L.;
RT   "Purification and immunological analysis of phospholipase D from
RT   castor bean endosperm.";
RL   Arch. Biochem. Biophys. 306:486-494(1993).
DR   GO; GO:0004630; F:phospholipase D activity; IEA.
SQ   SEQUENCE   15 AA;  1620 MW;  7C3849DA9BB2F50C CRC64;

  Query Match             36.8%;  Score 28;  DB 2;  Length 15;
  Best Local Similarity   83.3%;  Pred. No. 8.1e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 KTVGFG 14
              :|||||
Db          8 ETVGFG 13


RESULT 4
P70007
ID   P70007      PRELIMINARY;      PRT;    14 AA.
AC   P70007;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Histone H4-1 (Fragment).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=85160855; PubMed=4039007;
RA   Gargiulo G., Razvi F., Ruberti I., Mohr I., Worcel A.;
RT   "Chromatin-specific hypersensitive sites are assembled on a Xenopus
RT   histone gene injected into Xenopus oocytes.";
RL   J. Mol. Biol. 181:333-349(1985).
DR   EMBL; M23777; AAA49737.1; -.
DR   PIR; I51432; I51432.
FT   NON_TER       1      1
FT   CHAIN        <1     14       Potential.
SQ   SEQUENCE   14 AA;  1524 MW;  65A76B0A927B34B4 CRC64;

  Query Match             31.6%;  Score 24;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 GFGG 15
              ||||
Db         11 GFGG 14


RESULT 5
Q8TCS7
ID   Q8TCS7      PRELIMINARY;      PRT;    15 AA.
AC   Q8TCS7;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Complement component C7 (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Gonzalez S., Martinez Borra J., Lopez Larrea C.;
RT   "Cloning and characterization of human complement component C7
RT   promoter.";
RL   Genes Immun. 0:0-0(2002).
DR   EMBL; Y11720; CAA72407.1; -.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1666 MW;  10984ADCA7B43F19 CRC64;

  Query Match             31.6%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 4.3e+03;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 LFKTVGFGG 15
              ||  ||| |
Db          6 LFILVGFIG 14


RESULT 6
Q46963
ID   Q46963      PRELIMINARY;      PRT;    15 AA.
AC   Q46963;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Spot 42 RNA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CP78;
RX   MEDLINE=80074983; PubMed=390161;
RA   Sahagan B.G., Dahlberg J.E.;
RT   "A small, unstable RNA molecule of Escherichia coli: spot 42 RNA.";
RL   J. Mol. Biol. 131:573-592(1979).
DR   EMBL; X01895; CAA25985.1; -.
DR   PIR; A26228; A26228.
SQ   SEQUENCE   15 AA;  1725 MW;  5A51DAF170EA661E CRC64;

  Query Match             31.6%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 4.3e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 LFKTVGFG 14
              |   :|||
Db          8 LLHVIGFG 15


RESULT 7
Q7CPD2
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-14
Perfect score: 76
Sequence: 1 MARAAFLFKTVGFGG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	70	92.1	14	2	Q85718	Q85718 reovirus ty
2	29.5	38.8	14	2	Q85662	Q85662 reovirus ty
3	28	36.8	15	2	Q9S8Q8	Q9s8q8 ricinus com
4	24	31.6	14	2	P70007	P70007 xenopus lae
5	24	31.6	15	2	Q8TCS7	Q8tcs7 homo sapien
6	24	31.6	15	2	Q46963	Q46963 escherichia
7	24	31.6	15	2	Q7CPD2	Q7cpd2 salmonella
8	23	30.3	10	2	Q70F01	Q70f01 sus scrofa
9	23	30.3	10	2	Q9TG68	Q9tg68 gerrhonotus
10	23	30.3	15	2	Q7M4Z7	Q7m4z7 fusarium sp
11	22	28.9	10	1	CU30_LOCMI	P11735 locusta mig
12	22	28.9	10	2	Q85J75	Q85j75 varanus bre
13	22	28.9	10	2	Q9TG32	Q9tg32 ophisaurus
14	22	28.9	10	2	Q9TG47	Q9tg47 ophisaurus
15	22	28.9	11	2	Q94VH7	Q94vh7 varanus gil
16	22	28.9	13	2	O99783	O99783 caprimulgus
17	22	28.9	15	2	Q8SJ19	Q8sj19 phalacrocor
18	22	28.9	15	2	Q6JCR8	Q6jcr8 bemisia arg
19	22	28.9	15	2	Q6QQR3	Q6qqr3 helicobacte
20	21	27.6	10	1	ODP2_BOVIN	P11180 bos taurus
21	21	27.6	10	2	Q70LS6	Q70ls6 homo sapien
22	21	27.6	11	2	Q9GH12	Q9gh12 pandorina m
23	21	27.6	11	2	Q86864	Q86864 lymphocytic
24	21	27.6	12	2	Q6AV52	Q6av52 oryza sativ
25	21	27.6	15	2	Q7M4Z8	Q7m4z8 fusarium sp
26	20	26.3	10	2	Q6UJI5	Q6uji5 nephrurus w
27	20	26.3	11	2	Q86866	Q86866 lymphocytic
28	20	26.3	11	2	Q86868	Q86868 lymphocytic
29	20	26.3	12	2	Q841R5	Q841r5 agrobacteri
30	20	26.3	12	2	Q91YF5	Q91yf5 mus musculu
31	20	26.3	13	2	Q7Z3E0	Q7z3e0 homo sapien
32	20	26.3	13	2	O97122	O97122 toxoplasma
33	20	26.3	13	2	Q9U7D6	Q9u7d6 neospora ca
34	20	26.3	14	1	GLPK_STRGR	P25013 streptomyce
35	20	26.3	15	2	Q94IJ1	Q94ij1 zea mays (m
36	20	26.3	15	2	Q7M030	Q7m030 rattus norv
37	20	26.3	15	2	Q80X05	Q80x05 mesocricetu
38	19	25.0	10	2	Q7M3T6	Q7m3t6 tripneustes
39	19	25.0	10	2	O79924	O79924 elgaria pan
40	19	25.0	10	2	Q6UJJ4	Q6ujj4 nephrurus l
41	19	25.0	10	2	Q6UJJ7	Q6ujj7 carphodacty
42	19	25.0	10	2	Q6UJK6	Q6ujk6 rhynchoedur
43	19	25.0	10	2	Q6UJP5	Q6ujp5 diplodactyl
44	19	25.0	10	2	Q9TFV5	Q9tfv5 eublepharus
45	19	25.0	10	2	Q9TG50	Q9tg50 elgaria mul

ALIGNMENTS

RESULT 1
Q85718
ID Q85718 PRELIMINARY; PRT; 14 AA.
AC Q85718;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Reovirus serotype 1 S2 (Fragment).
OS Reovirus type 1 (strain Lang) (T1L) (Mammalian orthoreovirus 1).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus;
OC Mammalian orthoreoviruses.
OX NCBI_TaxID=10884;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83017877; PubMed=7123853;
RA Gaillard R.K., Li J.K., Keene J.D., Joklik W.K.;
RT "The sequences at the termini of four genes of the three reovirus
RT serotypes.";
RL Virology 121:320-326(1982).
DR EMBL; J02303; AAA47241.1; -.
DR InterPro; IPR004317; Sigma_1_2.
DR Pfam; PF03084; Sigma_1_2; 1.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1516 MW; 52F79D201BF900C7 CRC64;

Query Match 92.1%; Score 70; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MARAAFLFKTVGFG 14
||||||||||||||
Db 1 MARAAFLFKTVGFG 14

RESULT 2
Q85662
ID Q85662 PRELIMINARY; PRT; 14 AA.
AC Q85662;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Reovirus serotype 2 S2 (Fragment).
OS Reovirus type 2 (strain D5/Jones) (T2J) (Mammalian orthoreovirus 2).
OC Viruses; dsRNA viruses; Reoviridae; Orthoreovirus;
OC Mammalian orthoreoviruses.
OX NCBI_TaxID=10885;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83017877; PubMed=7123853;
RA Gaillard R.K., Li J.K., Keene J.D., Joklik W.K.;
RT "The sequences at the termini of four genes of the three reovirus
RT serotypes.";
RL Virology 121:320-326(1982).
DR EMBL; J02311; AAA47250.1; -.

```
Db          9 VGAGG 13


RESULT 14
PH0751
T-cell receptor beta chain (F12) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C;Accession: PH0751
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0751
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60843; NID:g50931; PIDN:CAA43235.1; PID:g50932
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             26.3%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 3.2e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         10 TVGFGG 15
              ::| ||
Db          3 SIGTGG 8


RESULT 15
C34874
transforming protein (N-rasB) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 09-Jul-2004
C;Accession: C34874
R;McMahon, G.; Davis, E.F.; Huber, L.J.; Kim, Y.; Wogan, G.N.
Proc. Natl. Acad. Sci. U.S.A. 87, 1104-1108, 1990
A;Title: Characterization of c-Ki-ras and N-ras oncogenes in aflatoxin B-1-induced rat l
A;Reference number: A34874; MUID:90138946; PMID:2105496
A;Accession: C34874
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <MCM>
A;Cross-references: UNIPROT:Q7M030

  Query Match             26.3%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 VGFGG 15
              || ||
Db          2 VGAGG 6


Search completed: February 22, 2005, 09:46:30
Job time : 12.1333 secs
```

```
  Best Local Similarity   75.0%;  Pred. No. 2.1e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 GFGG 15
              |:||
Db          2 GYGG 5


RESULT 9
PA0064
cytochrome C 1 - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0064
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0064
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Z8
C;Keywords: electron transfer; oxidative phosphorylation

  Query Match             27.6%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   83.3%;  Pred. No. 2.1e+03;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 AFLFKT 10
              | ||||
Db          9 ANLFKT 14


RESULT 10
S57574
T cell receptor V-J junctional alpha chain region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 17-Nov-1995 #text_change 05-Nov-1999
C;Accession: S57574
R;Burrows, S.R.; Silins, S.L.; Moss, D.J.; Khanna, R.; Misko, I.S.; Argaet, V.P.
submitted to the EMBL Data Library, June 1995
A;Description: T cell receptor repertoire for a viral epitope in humans is diversified b
A;Reference number: S57494
A;Accession: S57574
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-14 <BUR>
A;Cross-references: EMBL:Z49959; NID:g887504; PIDN:CAA90230.1; PID:g887505
C;Keywords: T-cell receptor

  Query Match             27.0%;  Score 20.5;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 2.4e+03;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    1;  Gaps    1;

Qy          8 FKTVGFG 14
              |||: ||
Db          9 FKTI-FG 14


RESULT 11
S26549
T-cell receptor beta chain (clone Cw3/A8, Cw3/Cas1) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C;Accession: S26549; S26550
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wid
J. Exp. Med. 176, 439-447, 1992
A;Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell receptor
A;Reference number: S26512; MUID:92364546; PMID:1380061
A;Accession: S26549
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Cross-references: EMBL:X67999
A;Experimental source: cytolytic T-lymphocyte, clone Cw3/A8
A;Accession: S26550
A;Molecule type: mRNA
A;Residues: 1-12 <CA2>
A;Cross-references: EMBL:X68000
A;Experimental source: cytolytic T-lymphocyte, clone Cw3/Cas1
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

  Query Match             26.3%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   45.5%;  Pred. No. 2.6e+03;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 AAFLFKTVGFG 14
              |: | :|: ||
Db          2 ASSLGETLYFG 12


RESULT 12
PQ0058
glycerol kinase (EC 2.7.1.30) - Streptomyces griseus (fragment)
C;Species: Streptomyces griseus
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
C;Accession: PQ0058
R;Bolotin, A.; Biro, S.
Gene 87, 151-152, 1990
A;Title: Nucleotide sequence of the putative regulatory gene and major promoter region c
A;Reference number: JQ0490; MUID:90236293; PMID:2110096
A;Accession: PQ0058
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-14 <BOL>
A;Cross-references: UNIPROT:P25013; GB:M37327; NID:g153287; PIDN:AAA26751.1; PID:g153289
C;Genetics:
A;Gene: glyA
A;Start codon: GTG
C;Keywords: phosphotransferase

  Query Match             26.3%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 3e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          6 FLFKTVG 12
              |: :|:|
Db          7 FIGETIG 13


RESULT 13
I52734
gene c-Ki-ras protein - hamster (fragment)
C;Species: Cricetinae gen. sp. (hamster)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 28-May-1999
C;Accession: I52734
R;Takahashi, T.; Moyer, M.P.; Cano, M.; Wang, Q.J.; Mountjoy, C.P.; Sanger, W.; Adrian,
Carcinogenesis 16, 931-939, 1995
A;Title: Differences in molecular biological, biological and growth characteristics betw
A;Reference number: I52734; MUID:95246257; PMID:7728976
A;Accession: I52734
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-15 <RES>
A;Cross-references: GB:S77068; NID:g914176
C;Genetics:
A;Gene: c-Ki-ras
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding

  Query Match             26.3%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 VGFGG 15
              || ||
```

```
spot 42 protein - Escherichia coli
C;Species: Escherichia coli
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: A26228; A37586; Q00551
R;Joyce, C.M.; Grindley, N.D.F.
J. Bacteriol. 152, 1211-1219, 1982
A;Title: Identification of two genes immediately downstream from the polA gene of Escher
A;Reference number: A26228; MUID:83056713; PMID:6183253
A;Accession: A26228
A;Molecule type: DNA
A;Residues: 1-15 <JOY2>
A;Cross-references: UNIPROT:Q46963
R;Sahagan, B.G.; Dahlberg, J.E.
J. Mol. Biol. 131, 573-592, 1979
A;Title: A small, unstable RNA molecule of Escherichia coli: spot 42 RNA.
A;Reference number: A37586; MUID:80074983; PMID:390161
A;Accession: A37586
A;Molecule type: DNA
A;Residues: 1-15 <SAH>
A;Cross-references: GB:X01895; NID:g40868; PIDN:CAA25985.1; PID:g40869
C;Genetics:
A;Gene: spf
A;Map position: 87 min

  Query Match             31.6%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 6e+02;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          7 LFKTVGFG 14
              |   :|||
Db          8 LLHVIGFG 15


RESULT 4
PA0087
cytochrome c2 - fungus (Fusarium sporotrichioides) (fragment)
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
C;Accession: PA0087
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Description: Two dimensional polyacrylamide gel electropheresis of Fusarium sporotrich
A;Reference number: PA0051
A;Accession: PA0087
A;Molecule type: protein
A;Residues: 1-15 <CHO>
A;Cross-references: UNIPROT:Q7M4Z7
C;Keywords: electron transfer; heme; photosynthesis

  Query Match             30.3%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 9.1e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 RAAFLFKT 10
              : | ||||
Db          7 KGANLFKT 14


RESULT 5
A58375
microcin B17 - Escherichia coli (fragment)
C;Species: Escherichia coli plasmid pPY113
C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
C;Accession: A58375
R;Yorgey, P.; Lee, J.; Koerdel, J.; Vivas, E.; Warner, P.; Jebaratnam, D.; Kolter, R.
Proc. Natl. Acad. Sci. U.S.A. 91, 4519-4523, 1994
A;Title: Posttranslational modifications in microcin B17 define an additional class of D
A;Reference number: A58375; MUID:94240167; PMID:8183941
A;Accession: A58375
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <YOR>
A;Cross-references: UNIPROT:P05834

  Query Match             28.9%;  Score 22;  DB 2;  Length 12;
  Best Local Similarity   80.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         11 VGFGG 15
              || ||
Db          1 VGIGG 5


RESULT 6
PT0664
T-cell receptor beta chain V-D-J region (121-2L) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0664
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0664
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-10 <FBE>
A;Experimental source: day 4 postnatal thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             27.6%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         10 TVGFGG 15
              | | ||
Db          5 TPGLGG 10


RESULT 7
E33098
214K exoantigen (version 2) - malaria parasite (Plasmodium falciparum) (fragments)
C;Species: Plasmodium falciparum
C;Date: 24-Aug-1990 #sequence_revision 24-Aug-1990 #text_change 09-Jun-2000
C;Accession: E33098
R;Nichols, J.H.; Hager, L.P.
submitted to the Protein Sequence Database, May 1990
A;Reference number: A33098
A;Accession: E33098
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <NIC>

  Query Match             27.6%;  Score 21;  DB 2;  Length 14;
  Best Local Similarity   75.0%;  Pred. No. 2e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 GFGG 15
              |:||
Db          1 GYGG 4


RESULT 8
PC4269
unidentified QR310003 protein - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 28-May-1997 #sequence_revision 18-Jul-1997 #text_change 18-Jul-1997
C;Accession: PC4269
R;Kawakami, T.; Kamo, M.; Chen, M.C.; Tsugita, A.
submitted to JIPID, April 1997
A;Reference number: PC4267
A;Accession: PC4269
A;Molecule type: protein
A;Residues: 1-15 <KAW>

  Query Match             27.6%;  Score 21;  DB 2;  Length 15;
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-14
Perfect score: 76
Sequence: 1 MARAAFLPKTVGPGG 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	38.2	14	2	A17150	glucose 1-dehydrog
2	24	31.6	14	2	I51432	histone H4-1 precu
3	24	31.6	15	2	A26228	spot 42 protein -
4	23	30.3	15	2	PA0087	cytochrome c2 - fu
5	22	28.9	12	2	A58375	microcin B17 - Esc
6	21	27.6	10	2	PT0664	T-cell receptor be
7	21	27.6	14	2	E33098	214K exoantigen (v
8	21	27.6	15	2	PC4269	unidentified QR310
9	21	27.6	15	2	PA0064	cytochrome C 1 - f
10	20.5	27.0	14	2	S57574	T cell receptor V-
11	20	26.3	12	2	S26549	T-cell receptor be
12	20	26.3	14	2	PQ0058	glycerol kinase (E
13	20	26.3	15	2	I52734	gene c-Ki-ras prot
14	20	26.3	15	2	PH0751	T-cell receptor be
15	20	26.3	15	2	C34874	transforming prote
16	19	25.0	10	2	C60527	sperm-activating p
17	19	25.0	11	2	PT0301	Ig heavy chain CRD
18	19	25.0	11	2	C58501	42K bile stone pro
19	19	25.0	11	2	D42965	talin - chicken (f
20	19	25.0	12	2	PH1463	T-cell receptor be
21	19	25.0	13	2	E42762	proteasome endopep
22	19	25.0	14	2	B61597	cytochrome P450 AL
23	19	25.0	15	2	S03353	plastocyanin - Mic
24	18	23.7	4	2	A53284	T-cell receptor be
25	18	23.7	7	2	H33098	180K exoantigen -
26	18	23.7	7	2	PT0663	T-cell receptor be
27	18	23.7	8	2	PT0725	T-cell receptor be
28	18	23.7	8	2	PC4373	telomeric and tetr
29	18	23.7	8	2	B24749	neuropeptide B - b
30	18	23.7	9	2	D24180	fibrinogen beta ch
31	18	23.7	9	2	F28854	fibrinopeptide B -
32	18	23.7	9	2	PT0225	Ig heavy chain CDR
33	18	23.7	9	2	S10784	enamelin i - bovin
34	18	23.7	9	2	PC7074	translation elonga
35	18	23.7	10	1	ECLQ4M	tachykinin IV - mi
36	18	23.7	10	2	PT0632	T-cell receptor be
37	18	23.7	10	2	S06964	hypothetical prote
38	18	23.7	10	2	I40032	trpE protein - Bac
39	18	23.7	10	2	F60527	sperm-activating p
40	18	23.7	10	2	B60589	sperm-activating p
41	18	23.7	11	2	A44755	20alpha-hydroxyste
42	18	23.7	11	2	A38590	transforming prote
43	18	23.7	12	1	LFECPE	pyrE leader peptid
44	18	23.7	12	2	A49261	coagulation factor
45	18	23.7	13	2	S48210	collagen alpha 1(V

ALIGNMENTS

RESULT 1
A17150
glucose 1-dehydrogenase (NAD) (EC 1.1.1.118) - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 03-Jun-2002
C;Accession: A17150
R;Franzen, B.; Carrubba, C.; Feingold, D.S.; Ashcom, J.; Franzen, J.S.
Biochem. J. 199, 599-602, 1981
A;Title: Amino acid sequence of the tryptic peptide containing the catalytic-site thiol
A;Reference number: A17150; MUID:82182061; PMID:6896145
A;Accession: A17150
A;Molecule type: protein
A;Residues: 1-14 <FRA>
C;Keywords: NAD; oxidoreductase

Query Match 38.2%; Score 29; DB 2; Length 14;
Best Local Similarity 83.3%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy         10 TVGPGG 15
              :|||||
Db          2 SVGPGG 7
```

RESULT 2
I51432
histone H4-1 precursor - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51432
R;Gargiulo, G.; Razvi, F.; Ruberti, I.; Mohr, I.; Worcel, A.
J. Mol. Biol. 181, 333-349, 1985
A;Title: Chromatin-specific hypersensitive sites are assembled on a Xenopus histone gene
A;Reference number: I51431; MUID:85160855; PMID:4039007
A;Accession: I51432
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-14 <GAR>
A;Cross-references: UNIPROT:P70007; GB:M23777; NID:g214219; PIDN:AAA49737.1; PID:g214222
C;Superfamily: histone H4

Query Match 31.6%; Score 24; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         12 GFGG 15
              ||||
Db         11 GFGG 14
```

RESULT 3
A26228

DR WPI; 2003-903164/82.
XX
PT New recombinant bacterial enzymes involved in nucleotide transport and
PT metabolism, useful for designing potential antibacterial agents.
XX
PS Disclosure; Fig 49; 392pp; English.
XX
CC This invention relates to novel isolated, recombinant polypeptides that
CC are involved in nucleotide transport and metabolism. Specifically, it
CC refers to the identification and functional characterisation of enzymes
CC that are antimicrobial targets from pathogenic bacteria, namely
CC Staphylococcus aureus, Pseudomonas aeruginosa, Escherichia coli,
CC Enterococcus faecalis, Haemophilus influenzae and Streptococcus
CC pneumoniae. These enzymes are deoxyuridine 5' triphosphate
CC nucleotidohydrolase (dut), guanylate kinase (KGUA), adenine
CC phosphoribosyltransferase (APT), phosphoribosylpyrophosphate synthetase
CC (PRSA), thymidylate synthase (thyA), uridylate kinase (PYRH), ribose
CC phosphate pyrophosphokinase and cytidine/ deoxyctidylate deaminase family
CC protein (YHFC). The present invention describes modulators, inhibitors,
CC agonists and antagonists against the proteins, which can be used for the
CC diagnosis and treatment of various diseases and conditions specific to
CC each bacterial species. Accordingly, the compositions of this invention
CC can be used in the development of drugs to treat, for example, impetigo,
CC cystitis, meningitis, osteomyelitis, otitis media and bacteraemia.
CC Furthermore, determination of the crystalline structure can be useful for
CC the design of modulators as potential therapeutic agents. This peptide
CC sequence is an Enterococcus faecalis thyA peptide derived from a tryptic
CC mass spectrum of the invention.
XX
SQ Sequence 14 AA;

Query Match 37.9%; Score 33; DB 7; Length 14;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy          4 LVWFIKKNT 12
              |:||:| :|
Db          3 LLWFLKGDT 11
```

RESULT 15
AAR99504
ID AAR99504 standard; peptide; 10 AA.
XX
AC AAR99504;
XX
DT 09-JAN-1997 (first entry)
XX
DE B-cell epitope used in construction of chimeric immunoglobulin.
XX
KW Chimera; chimeric; immunoglobulin; Ig; B-cell; T-cell; lymphocyte;
KW epitope; immune response; vaccine; pathogen; antibody; influenza;
KW measles; hepatitis; foot and mouth disease; tetanus toxoid;
KW human immunodeficiency virus; HIV; heat shock protein; M protein;
KW hen egg white lysozyme; nuclease.
XX
OS Influenza virus.
XX
PN WO9619584-A1.
XX
PD 27-JUN-1996.
XX
PF 21-DEC-1995; 95WO-US016718.
XX
PR 22-DEC-1994; 94US-00363276.
XX
PA (MOUN) MOUNT SINAI SCHOOL MEDICINE.
XX
PI Bona C, Zaghouani H;
XX
DR WPI; 1996-309598/31.
XX
PT Chimeric immunoglobulin with CDR loop substd. for T and or B cell epitope
PT - useful in vaccine composition to enhance immune response to pathogens.
XX
PS Disclosure; Page 17; 131pp; English.
XX
CC Chimeric immunoglobulins (Ig) having a CDR loop of the parent Ig replaced
CC with a foreign peptide sequence corresponding to a T- or B- cell epitope,
CC may be used in vaccine compositions to enhance an immune response to a
CC pathogen. Chimeric Ig comprising a B-cell epitope can also be used to
CC label B-cells, to test the ability of a subject to mount a humoural
CC response to a particular B-cell epitope or to collect B-cells which
CC recognise the epitope. . An antibody comprising a chimeric Ig molecule
CC which comprises an antigen binding site may be used in diagnostic assays
CC to detect the presence of a particular target antigen, which binds to the
CC antibody binding site. Sequences of B-cell epitopes are given in AAR99503
CC -08. Sequences of T-cell epitopes are given in AAT99509-17. The term B-
CC cell epitope refers to a peptide which is able to bind to an
CC immunoglobulin receptor of a B-cell and participate in the induction of
CC antibody production by that B-cell. This peptide is an immunodominant B-
CC cell epitope in site B of influenza virus HA1 haemagglutinin
XX
SQ Sequence 10 AA;

Query Match 36.8%; Score 32; DB 2; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy          6 WFIKKNTRYP 15
              |  ||   ||
Db          1 WLTKKGDSYP 10
```

Search completed: February 22, 2005, 09:24:54
Job time : 67.6667 secs

PT New composition of peptides and nucleic acids capable of binding Major
PT Histocompatibility Complex molecules, useful for diagnosing, preventing
PT or treating viral infections or cancer, such as prostate cancer,
PT hepatitis B or AIDS.
XX
PS Claim 1; SEQ ID NO 2222; 186pp; English.
XX
CC The invention relates to a novel composition comprising one or more
CC peptides or nucleic acids encoding an HLA binding peptide. The
CC composition further comprises an HTL epitope. It also comprises a spacer
CC molecule, a carrier, an MHC targeting sequence or a lipid. The peptides
CC are incorporated as part of a liposome. The peptide is from an antigen
CC selected from prostate specific antigen (PSA), prostate specific membrane
CC antigen (PSM), hepatitis B virus (HBV) antigen, hepatitis C virus (HCV)
CC antigen, malignant melanoma antigen (MAGE), Epstein Barr virus, human
CC immunodeficiency type-1 (HIV-I), human immunodeficiency type-2 (HIV-2),
CC Papilloma virus, Lassa virus, Mycobacterium tuberculosis (MT), p53,
CC murine p53 (mp53), CEA, HER2/neu, and tyrosine kinase related protein
CC (TKP). The composition is useful for preventing or treating viral
CC infections or cancer, such as prostate cancer, hepatitis B, hepatitis C,
CC AIDS, renal carcinoma, cervical carcinoma, lymphoma, CMV or chondyloma
CC acuminatum. The composition is also be used for diagnosing such diseases.
CC This sequence represents a peptide of the invention.
XX
SQ Sequence 9 AA;

```
  Query Match             39.1%;  Score 34;  DB 8;  Length 9;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YRNLVW 6
              ||||:|
Db          3 YRNLLW 8
```

RESULT 13
ADR99836
ID ADR99836 standard; peptide; 13 AA.
XX
AC ADR99836;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human aggrecan peptide P4R as intermediate for citrullination method.
XX
KW immunosuppressive; antiarthritic; antirheumatic; neuroprotective;
KW antiinflammatory; gene therapy; peptidyl arginine deaminase inhibitor;
KW citrullinated peptide; major histocompatibility class II molecule;
KW epitope; myelin basic protein; MBP; glial fibrillary acid protein; GFAP;
KW diagnosis; autoimmune disorder; rheumatoid arthritis; multiple sclerosis.
XX
OS Homo sapiens.
XX
PN WO2004078098-A2.
XX
PD 16-SEP-2004.
XX
PF 05-MAR-2004; 2004WO-CA000337.
XX
PR 07-MAR-2003; 2003US-0452522P.
XX
PA (LONH-) LONDON HEALTH SCI CENT RES INC.
XX
PI Hill J, Cairns E, Bell D;
XX
DR WPI; 2004-668479/65.
XX
PT New citrullinated myelin basic protein or glial fibrillary acid protein
PT peptide, for use in preparing a composition for diagnosing or treating
PT rheumatoid arthritis or multiple sclerosis.
XX
PS Example 2; Page 55; 93pp; English.
XX
CC The invention relates to a citrullinated peptide, binding with increased
CC affinity to a major histocompatibility (MHC) class II molecule having a
CC shared epitope comprising an amino acid sequence selected from ADR99775-
CC ADR99811, given in the specification. The citrullinated myelin basic
CC protein (MBP) and/or glial fibrillary acid protein (GFAP) peptide is
CC useful in preparing a composition for diagnosing or treating autoimmune
CC disorder, e.g., rheumatoid arthritis or multiple sclerosis. This sequence
CC represents a fragment of the human aggrecan protein (aa 280-292) in which
CC the Asp residue at position 6 is altered to an Arg residue and used to
CC generate a citrullinated peptide of the invention.
XX
SQ Sequence 13 AA;

```
  Query Match             39.1%;  Score 34;  DB 8;  Length 13;
  Best Local Similarity   40.0%;  Pred. No. 85;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 WFIKKNTRYP 15
              |  :::  |||
Db          3 WLARRSVRYP 12
```

RESULT 14
ADJ62411
ID ADJ62411 standard; peptide; 14 AA.
XX
AC ADJ62411;
XX
DT 06-MAY-2004 (first entry)
XX
DE Tryptic mass spectrum E_faecalis thyA peptide 11.
XX
KW pathogenic bacteria; deoxyuridine 5' triphosphate nucleotidohydrolase;
KW dut; guanylate kinase; KGUA; adenine phosphoribosyltransferase; APT;
KW phosphoribosylpyrophosphate synthetase; PRSA; thymidylate synthase; thyA;
KW KTHY; uridylate kinase; PYRH; ribose phosphate pyrophosphokinase;
KW cytidine/ deoxyctidylate deaminase; YHFC; impetigo; cystitis; meningitis;
KW osteomyelitis; otitis media; bacteraemia; nucleotide transport;
KW antimicrobial.
XX
OS Enterococcus faecalis.
XX
PN WO2003087354-A2.
XX
PD 23-OCT-2003.
XX
PF 08-APR-2003; 2003WO-CA000485.
XX
PR 09-APR-2002; 2002US-0371067P.
PR 05-JUN-2002; 2002US-0386548P.
PR 06-JUN-2002; 2002US-0386826P.
PR 06-JUN-2002; 2002US-0386869P.
PR 06-NOV-2002; 2002US-0424380P.
PR 08-NOV-2002; 2002US-0425086P.
PR 24-DEC-2002; 2002US-0436243P.
PR 24-DEC-2002; 2002US-0436288P.
PR 26-DEC-2002; 2002US-0436566P.
PR 26-DEC-2002; 2002US-0436567P.
PR 27-DEC-2002; 2002US-0436708P.
PR 30-DEC-2002; 2002US-0436947P.
PR 30-DEC-2002; 2002US-0436971P.
PR 30-DEC-2002; 2002US-0437038P.
PR 30-DEC-2002; 2002US-0437141P.
PR 31-DEC-2002; 2002US-0437620P.
PR 31-DEC-2002; 2002US-0437638P.
XX
PA (AFFI-) AFFINIUM PHARM INC.
XX
PI Edwards A, Dharamsi A, Vedadi M, Domagala M, Mansoury K;
PI Houston S, Awrey D, Beattie B, Kanagarajah D, Vallee F, Virag C;
PI Buzadzija K, Mcdonald M, Tai M, Pinder B, Alam MZ, Ouyang H;
PI Richards D, Canadien V, Thalakada R, Nethery K;
XX

```
PA   (KERN/) KERN F.
XX
PI   Kern F,  Volk H,  Reinke P,  Faulhaber N,  Surel I,  Khatamzas E;
XX
DR   WPI; 2001-557718/62.
XX
PT   Stimulating CD8 or CD4 T cells for the immuno-stimulation of mammals and
PT   the determination of the response to the antigen comprises using a
PT   synthetic peptide library of the specific antigen.
XX
PS   Example; Fig 1; 85pp; German.
XX
CC   This sequence represents a novel method for stimulating CD8 or CD4 T
CC   cells for the immuno-stimulation of mammals and the determination of the
CC   response to an antigen (I). The method comprises (i) dividing the amino
CC   acid sequence of the antigen into protein fragments (II) of at least 9
CC   amino acids, whereby adjacent or neighbouring fragments are in the whole
CC   antigen sequence; (ii) synthesizing a peptide library containing (II);
CC   and (iii) incubating the CD8 and/or CD4 T lymphocytes in a suspension
CC   comprising (II) in a single culture vessel. The method is used to immuno-
CC   stimulate T cells of mammals, particularly humans, and is also useful for
CC   diagnostic purposes to determine if a mammal, especially human, has
CC   previously immunologically responded to a specific protein, and to
CC   determine strength of that response. The method is suitable for in vivo
CC   or in vitro immuno-stimulation of mammalian, more preferably human T
CC   lymphocytes, where the stimulated T lymphocytes are expanded and can be
CC   transfused into a patient. AAB86544-AAB86803 represent fragments of the
CC   human cytomegalovirus IE1 and pp65 lower matrix phosphoprotein which are
CC   used to illustrate the method of the invention
XX
SQ   Sequence 15 AA;

  Query Match             39.7%;  Score 34.5;  DB 4;  Length 15;
  Best Local Similarity   46.7%;  Pred. No. 81;
  Matches    7;  Conservative    4;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 YRNLVWPIKKNTRYP 15
              |||: :|  ||: :|
Db          1 YRNIEFF-TKNSAFP 14


RESULT 11
ADN65655
ID   ADN65655 standard; peptide; 9 AA.
XX
AC   ADN65655;
XX
DT   01-JUL-2004  (first entry)
XX
DE   HLA binding peptide #2255.
XX
KW   cytostatic; hepatotropic; virucide; antiinflammatory; anti-HIV;
KW   gene therapy; vaccine; HLA binding peptide; HTL epitope; liposome;
KW   prostate specific antigen; prostate specific membrane antigen;
KW   hepatitis B virus antigen; hepatitis C virus antigen;
KW   malignant melanoma antigen; MAGE; Epstein Barr virus; cancer;
KW   prostate cancer; AIDS; renal carcinoma; cervical carcinoma; lymphoma;
KW   chondyloma acuminatum.
XX
OS   Unidentified.
XX
PN   WO2004031211-A2.
XX
PD   15-APR-2004.
XX
PF   03-OCT-2003; 2003WO-US031308.
XX
PR   03-OCT-2002; 2002US-0416207P.
PR   08-OCT-2002; 2002US-0417269P.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sidney J,  Southwood S,  Sette A;
XX
DR   WPI; 2004-347953/32.
XX
PT   New composition of peptides and nucleic acids capable of binding Major
PT   Histocompatibility Complex molecules, useful for diagnosing, preventing
PT   or treating viral infections or cancer, such as prostate cancer,
PT   hepatitis B or AIDS.
XX
PS   Claim 1; SEQ ID NO 2255; 186pp; English.
XX
CC   The invention relates to a novel composition comprising one or more
CC   peptides or nucleic acids encoding an HLA binding peptide. The
CC   composition further comprises an HTL epitope. It also comprises a spacer
CC   molecule, a carrier, an MHC targeting sequence or a lipid. The peptides
CC   are incorporated as part of a liposome. The peptide is from an antigen
CC   selected from prostate specific antigen (PSA), prostate specific membrane
CC   antigen (PSM), hepatitis B virus (HBV) antigen, hepatitis C virus (HCV)
CC   antigen, malignant melanoma antigen (MAGE), Epstein Barr virus, human
CC   immunodeficiency type-1 (HIV-I), human immunodeficiency type-2 (HIV-2),
CC   Papilloma virus, Lassa virus, Mycobacterium tuberculosis (MT), p53,
CC   murine p53 (mp53), CEA, HER2/neu, and tyrosine kinase related protein
CC   (TKP). The composition is useful for preventing or treating viral
CC   infections or cancer, such as prostate cancer, hepatitis B, hepatitis C,
CC   AIDS, renal carcinoma, cervical carcinoma, lymphoma, CMV or chondyloma
CC   acuminatum. The composition is also be used for diagnosing such diseases.
CC   This sequence represents a peptide of the invention.
XX
SQ   Sequence 9 AA;

  Query Match             39.1%;  Score 34;  DB 8;  Length 9;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+06;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YRNLVW 6
              ||||:|
Db          3 YRNLLW 8


RESULT 12
ADN65622
ID   ADN65622 standard; peptide; 9 AA.
XX
AC   ADN65622;
XX
DT   01-JUL-2004  (first entry)
XX
DE   HLA binding peptide #2222.
XX
KW   cytostatic; hepatotropic; virucide; antiinflammatory; anti-HIV;
KW   gene therapy; vaccine; HLA binding peptide; HTL epitope; liposome;
KW   prostate specific antigen; prostate specific membrane antigen;
KW   hepatitis B virus antigen; hepatitis C virus antigen;
KW   malignant melanoma antigen; MAGE; Epstein Barr virus; cancer;
KW   prostate cancer; AIDS; renal carcinoma; cervical carcinoma; lymphoma;
KW   chondyloma acuminatum.
XX
OS   Unidentified.
XX
PN   WO2004031211-A2.
XX
PD   15-APR-2004.
XX
PF   03-OCT-2003; 2003WO-US031308.
XX
PR   03-OCT-2002; 2002US-0416207P.
PR   08-OCT-2002; 2002US-0417269P.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sidney J,  Southwood S,  Sette A;
XX
DR   WPI; 2004-347953/32.
XX
```

PR 21-FEB-2002; 2002US-00082014.
PR 21-FEB-2003; 2003US-00372076.
XX
PA (PAGE/) PAGE M.
PA (FRIE/) FRIEDE M.
PA (SCHM/) SCHMIDT A E.
PA (STOB/) STOBER D.
XX
PI Page M, Friede M, Schmidt AE, Stober D;
XX
DR WPI; 2004-603322/58.
XX
PT Treating chronic hepatitis, by administering vaccine comprising
PT immunogenic particles having recombinant hepatitis B core chimeric
PT protein molecules, that stimulates T cell, to patient chronically
PT infected with hepatitis B virus.
XX
PS Disclosure; SEQ ID NO 14; 117pp; English.
XX
CC The invention relates to treating chronic hepatitis, by administering a
CC vaccine comprising immunogenic particles having recombinant hepatitis B
CC core (HBc) chimeric protein molecules (where truncated Hbc molecules are
CC linked N-terminally or C-terminally to an immunogenic epitope), that
CC stimulate T cell production, to a patient chronically infected with
CC hepatitis B virus, and maintaining patient for time sufficient to induce
CC T cells activated against HBc. The chimeric proteins are still capable
CC self-assembling into particles upon expression in a host cell and are
CC substantially free of binding to nucleic acids, and the particles display
CC enhanced stability. Also included is enhancing (M2) the production of one
CC or more of gamma-producing CD8+, CD4+ T cells and cytotoxic T lymphocytes
CC against hepatitis B virus, involving administering to a patient
CC chronically infected with hepatitis B virus, a T cell-stimulating amount
CC of a vaccine comprising immunogenic particles dissolved or dispersed in a
CC diluent containing one or both of an agonist of toll-like receptor 4 and
CC receptor 9 (TLR-4 and TLR-9), the immunogenic particles comprising HBc
CC chimeric protein molecules and maintaining the patient for a sufficient
CC time to induce T cells activated against HBc. The immunogenic epitopes
CC may be B cell or T cell epitopes. The chimeric vaccine is useful for
CC treating a patient chronically infected with hepatitis B virus. The
CC present sequence is a B cell epitope suitable for inclusion in the
CC chimeric protein of the invention.
XX
SQ Sequence 10 AA;

Query Match 44.8%; Score 39; DB 8; Length 10;
Best Local Similarity 60.0%; Pred. No. 9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YRNLVWFIKK 10
||||:| :|
Db 1 YRNLLWLTEK 10

RESULT 9
AAM97704
ID AAM97704 standard; peptide; 14 AA.
XX
AC AAM97704;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human peptide #979 encoded by a SNP oligonucleotide.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US035498.
XX
PR 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Disclosure; Page 3882; 4143pp; English.
XX
CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms
XX
SQ Sequence 14 AA;

Query Match 41.4%; Score 36; DB 4; Length 14;
Best Local Similarity 66.7%; Pred. No. 42;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 NLVWFIKKN 11
| || |:||
Db 6 NAVWLIQKN 14

RESULT 10
AAB86605
ID AAB86605 standard; peptide; 15 AA.
XX
AC AAB86605;
XX
DT 20-NOV-2001 (first entry)
XX
DE Human cytomegalovirus strain AD169 IE1 peptide fragment SEQ ID 62.
XX
KW Antigen-specific stimulation; T-lymphocyte; CD8 stimulation; pp65;
KW CD4 stimulation; immuno-stimulation; IE1; lower matrix phosphoprotein.
XX
OS Human cytomegalovirus.
XX
PN WO200163286-A2.
XX
PD 30-AUG-2001.
XX
PF 17-FEB-2001; 2001WO-EP001773.
XX
PR 22-FEB-2000; 2000DE-01009341.
XX

XX
DE Recombinant chimer hepatitis B core protein immunogenic epitope #6.
XX
KW Recombinant chimer hepatitis B core protein; HBc; immunogenic epitope;
KW HBc immunodominant loop; immune response.
XX
OS Influenza A virus.
XX
PN US2003185858-A1.
XX
PD 02-OCT-2003.
XX
PF 21-FEB-2002; 2002US-00082014.
XX
PR 15-AUG-2001; 2001US-00930915.
XX
PA (BIRK/) BIRKETT A J.
XX
PI Birkett AJ;
XX
DR WPI; 2004-031988/03.
XX
PT Recombinant chimer hepatitis B core protein molecule useful for preparing
PT vaccine or inoculum includes peptide-bonded heterologous immunogenic
PT epitope at N-terminus in the hepatitis B core immunodominant loop or C-
PT terminus of the chimer.
XX
PS Disclosure; SEQ ID NO 14; 110pp; English.
XX
CC The invention relates to a recombinant chimer hepatitis B core (HBc)
CC protein molecule that includes a peptide-bonded heterologous immunogenic
CC epitope at the N-terminus in the HBc immunodominant loop or the C-
CC terminus of the chimer, or a heterologous linker residue for a conjugated
CC epitope present in the loop. The invention also relates to an immunogenic
CC particle comprising the recombinant hepatitis B core chimeric protein
CC molecules, a vaccine comprising the immunogenic particles dissolved or
CC dispersed in a diluent, a nucleic acid that encodes a recombinant HBc
CC protein molecule or its variant, analogue, or complement and a method for
CC inducing an immune response in an inoculated host animal comprising
CC inoculating a host animal with a vaccine and maintaining the inoculated
CC animal for a period of time sufficient to enable development of an immune
CC response. The recombinant chimer hepatitis B core protein molecule is
CC used in an immunogenic particle for preparing a vaccine useful for
CC inducing an immune response in an inoculated host animal. This sequence
CC represents an HBc protein immunogenic B cell epitope of the invention.
XX
SQ Sequence 10 AA;

Query Match 44.8%; Score 39; DB 8; Length 10;
Best Local Similarity 60.0%; Pred. No. 9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YRNLVWFIKK 10
||||:| :|
Db 1 YRNLLWLTEK 10

RESULT 7
ADP73403
ID ADP73403 standard; peptide; 10 AA.
XX
AC ADP73403;
XX
DT 09-SEP-2004 (first entry)
XX
DE Influenza virus A8/PR8 B cell epitope of gene HA.
XX
KW transgenic animal; Hepatitis B virus nucleocaspid core protein; HBc;
KW enhanced stability; hepatotropic; virucide; immunology;
KW protein engineering; immunogen; vaccine; Hepatitis B infection.
XX
OS Influenza virus.
XX
PN WO2004053091-A2.
XX
PD 24-JUN-2004.
XX
PF 10-DEC-2003; 2003WO-US039164.
XX
PR 10-DEC-2002; 2002US-0432123P.
XX
PA (APOV-) APOVIA INC.
XX
PI Lyons K, Birkett AJ, Haron JA;
XX
DR WPI; 2004-468859/44.
XX
PT New recombinant chimer hepatitis B core (HBc) protein molecules useful in
PT the fields of immunology and protein engineering, in particular as an
PT immunogen in a vaccine for Hepatitis B infections.
XX
PS Disclosure; SEQ ID NO 16; 338pp; English.
XX
CC The invention relates to a novel recombinant chimeric Hepatitis B virus
CC nucleocaspid (core) protein (HBc), up to 600 or 380 amino acid residues
CC in length. The chimeric protein is engineered for both enhanced stability
CC of self-assembled particles and the substantial absence of nucleic acid
CC binding by the particles. The invention further comprises: a recombinant
CC HBc protein chimeric molecule that has a length of 135-365 amino acid
CC residues and contains four peptide-linked amino acid residue sequence
CC domains from the N-terminus that are denominated Domains I, II, III and
CC IV. The invention also provides nucleic acids, polypeptides, host cells,
CC vectors and transgenic animals used in the methods of the invention. The
CC chimeric compositions of the invention have hepatotropic and virucide
CC activities. The methods and compositions of the present invention are
CC useful in the fields of immunology and protein engineering, in particular
CC for using a chimeric hepatitis B virus nucleocapsid protein as an
CC immunogen in a vaccine for Hepatitis B infections. This sequence
CC represents a Hepatitis B virus nucleocaspid (core) protein related
CC polypeptide of the invention.
XX
SQ Sequence 10 AA;

Query Match 44.8%; Score 39; DB 8; Length 10;
Best Local Similarity 60.0%; Pred. No. 9;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YRNLVWFIKK 10
||||:| :|
Db 1 YRNLLWLTEK 10

RESULT 8
ADR12707
ID ADR12707 standard; peptide; 10 AA.
XX
AC ADR12707;
XX
DT 04-NOV-2004 (first entry)
XX
DE Influenza virus HA B cell epitope.
XX
KW HBV; chronic hepatitis; HBc; nucleocapsid core protein; vaccine;
KW immunogen; immunogenic epitope; T cell; B cell; CD8+ cell; CD4+ cell;
KW cytotoxic T lymphocyte; toll-like receptor 4; toll-like receptor 9;
KW TLR-4; TLR-9; epitope.
XX
OS Influenza virus; strain A8/PR8.
XX
PN US2004156863-A1.
XX
PD 12-AUG-2004.
XX
PF 01-OCT-2003; 2003US-00677074.
XX
PR 21-FEB-2002; 2002US-00080299.

```
  Query Match             44.8%;  Score 39;  DB 5;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 9;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YRNLVWFIKK 10
              ||||:|  :|
Db          1 YRNLLWLTEK 10


RESULT 4
ADE10780
ID   ADE10780 standard; peptide; 10 AA.
XX
AC   ADE10780;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Chimeric hepatitis B virus related B-cell epitope seqid 14.
XX
KW   hepatotropic; virucide; antiinflammatory; chronic hepatitis; vaccine;
KW   recombinant hepatitis B core chimeric protein; HBc chimeric protein;
KW   hepatitis B infection; T-cel stimulator; B-cell epitope.
XX
OS   Influenza virus.
XX
PN   US2003198645-A1.
XX
PD   23-OCT-2003.
XX
PF   21-FEB-2003; 2003US-00372076.
XX
PR   21-FEB-2002; 2002US-00080299.
PR   21-FEB-2002; 2002US-00082014.
XX
PA   (PAGE/) PAGE M.
PA   (FRIE/) FRIEDE M.
XX
PI   Page M,  Friede M;
XX
DR   WPI; 2003-852775/79.
XX
PT   Treating chronic hepatitis B infection by administering a T cell-
PT   stimulating vaccine containing immunogenic particles having recombinant
PT   carboxy-terminal truncated hepatitis B core (HBc) chimeric protein
PT   molecules.
XX
PS   Disclosure; SEQ ID NO 14; 111pp; English.
XX
CC   The invention describes a method of treating chronic hepatitis comprising
CC   administering to a patient a T cell-stimulating amount of a vaccine
CC   comprising immunogenic particles dissolved or dispersed in a diluent,
CC   where the immunogenic particles consists of recombinant hepatitis B core
CC   (HBc) chimeric protein molecules, and maintaining the patient to induce T
CC   cells activated against HBc. The methods and compositions of the present
CC   invention are useful for treating chromic hepatitis B infection. This is
CC   the amino acid sequence of a chimeric hepatitis B virus related B-cell
CC   epitope useful for expression within the HBV chimer at the N-terminus,
CC   within the immunoigenic loop and/or at the C-terminus.
XX
SQ   Sequence 10 AA;

  Query Match             44.8%;  Score 39;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 9;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YRNLVWFIKK 10
              ||||:|  :|
Db          1 YRNLLWLTEK 10


RESULT 5
ADM39685
ID   ADM39685 standard; peptide; 10 AA.
XX
AC   ADM39685;
XX
DT   03-JUN-2004  (first entry)
XX
DE   Influenza A virus B-cell peptide epitope expressed by HBc chimer Seq 17.
XX
KW   immunogenic; avian hepatitis B virus; nucleocapsid;
KW   self assembled particle; immunogen; inoculum; vaccine; immunostimulant;
KW   antibacterial; virucidal; B-cell epitope.
XX
OS   Influenza A virus.
XX
PN   WO2003072722-A2.
XX
PD   04-SEP-2003.
XX
PF   21-FEB-2003; 2003WO-US005315.
XX
PR   21-FEB-2002; 2002US-0359129P.
XX
PA   (APOV-) APOVIA INC.
XX
PI   Birkett AJ,  Peck B;
XX
DR   WPI; 2003-679948/64.
XX
PT   New recombinant chimer avian hepatitis B core protein molecule, useful as
PT   an immunogen for inducing a B cell or T cell response to produce
PT   antibodies, or as a vaccine against pathogens.
XX
PS   Disclosure; SEQ ID NO 17; 278pp; English.
XX
CC   This invention relates to novel recombinant immunogenic chimeric avian
CC   hepatitis B core (AHBc) nucleocapsid proteins. Specifically, it refers to
CC   an AHBc protein that has been engineered to display an immunogenic B cell
CC   or T cell epitope, exhibit enhanced stability and an absence of nucleic
CC   acid binding as a self assembled particle. The present invention
CC   describes the chimeric AHBc protein as truncated at the C-terminus and
CC   containing introduced cysteine residues that confers an enhanced
CC   stability in aqueous solution, an increased yield and more immunogenicity
CC   than similar conjugates that lack N- or C-terminal cysteines.
CC   Furthermore, a reduction in the number of positively charged residues
CC   (lysine and arginine) towards the C-terminus prepares self-assembled
CC   particles that are substantially free of nucleic acid binding. As such,
CC   these chimeric particles can be used as immunogens of an inoculum that
CC   induce a B cell or T cell response in an animal to produce antibodies. It
CC   can also be useful for developing a vaccine to protect against the
CC   pathogen from which the heterologous epitope or the hapten is derived.
CC   Accordingly, these compositions exhibit immunostimulant, antibacterial
CC   and virucidal activities. This peptide sequence is an exemplary B-cell
CC   epitope peptide immunogen useful for both linkage to the linker residue
CC   after expression of a contemplated chimer and for expression within an
CC   HBc chimer of the invention.
XX
SQ   Sequence 10 AA;

  Query Match             44.8%;  Score 39;  DB 7;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 9;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YRNLVWFIKK 10
              ||||:|  :|
Db          1 YRNLLWLTEK 10


RESULT 6
ADG63881
ID   ADG63881 standard; peptide; 10 AA.
XX
AC   ADG63881;
XX
DT   11-MAR-2004  (first entry)
```

```
CC   the autoimmune disease. This peptide is an internal peptide of influenza
CC   type A haemagglutinin protein and is implicated as a foreign epitope
CC   involved in the aetiology or in remissions of multiple sclerosis. It has
CC   been shown capable of inducing the proliferation of autoreactive T-cell
CC   clones isolated from multiple sclerosis patients. (Updated on 27-AUG-2003
CC   to correct OS field.)
XX
SQ   Sequence 15 AA;

  Query Match             100.0%;  Score 87;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 8.8e-08;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YRNLVWFIKKNTRYP 15
              |||||||||||||||
Db          1 YRNLVWFIKKNTRYP 15


RESULT 2
AAY29685
ID   AAY29685 standard; protein; 10 AA.
XX
AC   AAY29685;
XX
DT   08-NOV-1999  (first entry)
XX
DE   Influenza A8/PR8 antigen HA polypeptide hapten.
XX
KW   Human hepatitis B core protein; HBc; modified; immunodominant;
KW   nucleocapsid protein; vaccine; T cell epitope.
XX
OS   Influenza virus.
XX
PN   WO9940934-A1.
XX
PD   19-AUG-1999.
XX
PF   11-FEB-1999;   99WO-US003055.
XX
PR   12-FEB-1998;   98US-0074537P.
XX
PA   (IMMU-) IMMUNE COMPLEX CORP.
XX
PI   Birkett AJ;
XX
DR   WPI; 1999-527340/44.
XX
PT   Conjugate of hepatitis B core protein, modified to increase reactivity
PT   with hapten, used to raise antibodies against the hapten, e.g. in
PT   vaccines.
XX
PS   Example 3; Page 37; 128pp; English.
XX
CC   The present invention describes a conjugate (A) comprising a
CC   strategically modified hepatitis B core (HBc) protein (I) attached to a
CC   hapten, where (I) includes amino acids (aa) 10-140 of the wild type HBc
CC   183 aa sequence (given in AAY29674) and additionally has an insert (II)
CC   in the region corresponding to aa's 50-100, where the insert is of 1 to
CC   about 40 aa's and contains a chemically reactive aa residue linked to the
CC   hapten. A vaccine containing (A), optionally in the form of particles, is
CC   used to induce a protective antibody response against the pathogen from
CC   which the hapten is derived, in humans or other animals. These pathogens
CC   may be bacteria, viruses, rickettsia or protozoa. Insertion of (II)
CC   overcomes the low reactivity of aa side chains in native HBc protein,
CC   increasing the reactivity with hapten and resulting in conjugates of
CC   improved immunogenicity. Modified HBc can be derivatised in the form of
CC   particles by well-defined chemical methods, and is unlikely to cause
CC   immunological side-effects. AAY29675 to AAY29735 represent polypeptide
CC   haptens used in an example from the present invention
XX
SQ   Sequence 10 AA;

  Query Match             44.8%;  Score 39;  DB 2;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 9;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 YRNLVWFIKK 10
              ||||:|  :|
Db          1 YRNLLWLTEK 10


RESULT 3
AAU93809
ID   AAU93809 standard; peptide; 10 AA.
XX
AC   AAU93809;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Inlfuenza virus HA B cell epitope.
XX
KW   Immunogenic; hepadnavirus nucleocapsid protein; hepatitis B core; HBc;
KW   vaccine; B cell epitope; T cell epitope; immunostimulant.
XX
OS   Influenza virus.
XX
PN   WO200214478-A2.
XX
PD   21-FEB-2002.
XX
PF   16-AUG-2001; 2001WO-US041759.
XX
PR   16-AUG-2000; 2000US-0225843P.
PR   22-AUG-2000; 2000US-0226867P.
PR   15-AUG-2001; 2001US-00930915.
XX
PA   (APOV-) APOVIA INC.
XX
PI   Birkett AJ;
XX
DR   WPI; 2002-257601/30.
XX
PT   Novel recombinant hepadnavirus nucleocapsid protein, termed as chimeric
PT   hepatitis B core protein, displays immunogenic epitopes at N-terminus,
PT   HBc immunogenic loop with linker for conjugated epitope and C-terminus.
XX
PS   Disclosure; Page 35; 289pp; English.
XX
CC   The invention relates to a recombinant hepadnavirus nucleocapsid protein,
CC   i.e. a chimeric hepatitis B core (HBc) protein (I), displaying one or
CC   more immunogenic epitopes at the N-terminus, HBc immunogenic loop (L) or
CC   C-terminus, or having a heterologous linker for a conjugated epitope in
CC   (L), and containing a Cys residue at, or near, the C-terminus that
CC   confers enhanced stability to the particles. A vaccine comprising (I) is
CC   useful for inducing an immune response in an inoculated host animal, by
CC   inoculating a host animal with the vaccine, and maintaining that
CC   inoculated animal for a time period sufficient for that animal to develop
CC   an immune response. The immunogenic particles formed using (I) are
CC   substantially free of binding to nucleic acids, and are most stable than
CC   the particle formed from otherwise identical HBc chimer that lacks the C-
CC   terminal residue or in which a C-terminal Cys is replaced by another
CC   residue. The chimer particles are most stable on storage in aqueous
CC   compositions that are particles of similar sequence that lack any C-
CC   terminal Cys residues. The chimer molecule exhibits the self-assembly not
CC   exhibiting the nucleic acid binding of those native particles, and
CC   excellent B cell and T cell immunogenicities. The chimer particles are
CC   typically prepared in higher yield than similar particles that are free
CC   of a C-terminal Cys. The particles are often far more immunogenic than
CC   the similar conjugates that lack a C-terminal Cys. Immunogenicities of
CC   particles assembled from the chimer molecules are enhanced as compared to
CC   similar particles assembled from chimer molecules lacking at least one C-
CC   terminal Cys. AAU93802-AAU93997 represent immunogenic HBc particles amino
CC   acid sequences and related sequences of the invention
XX
SQ   Sequence 10 AA;
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:52:23 ; Search time 65.6667 Seconds
(without alignments)
88.346 Million cell updates/sec

Title: US-08-991-628-13
Perfect score: 87
Sequence: 1 YRNLVWFIKKNTRYP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 632537

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	87	100.0	15	2	AAW04853	Aaw04853 Internal
2	39	44.8	10	2	AAY29685	Aay29685 Influenza
3	39	44.8	10	5	AAU93809	Aau93809 Inlfuenza
4	39	44.8	10	7	ADE10780	Ade10780 Chimeric
5	39	44.8	10	7	ADM39685	Adm39685 Influenza
6	39	44.8	10	8	ADG63881	Adg63881 Recombina
7	39	44.8	10	8	ADP73403	Adp73403 Influenza
8	39	44.8	10	8	ADR12707	Adr12707 Influenza
9	36	41.4	14	4	AAM97704	Aam97704 Human pep
10	34.5	39.7	15	4	AAB86605	Aab86605 Human cyt
11	34	39.1	9	8	ADN65655	Adn65655 HLA bindi
12	34	39.1	9	8	ADN65622	Adn65622 HLA bindi
13	34	39.1	13	8	ADR99836	Adr99836 Human agg
14	33	37.9	14	7	ADJ62411	Adj62411 Tryptic m
15	32	36.8	10	2	AAR99504	Aar99504 B-cell ep
16	32	36.8	10	2	AAW12228	Aaw12228 Influenza
17	32	36.8	10	2	AAW60684	Aaw60684 Influenza
18	32	36.8	10	5	ABB79917	Abb79917 Influenza
19	31	35.6	13	2	AAW10481	Aaw10481 MHC Class
20	31	35.6	13	4	AAG65616	Aag65616 Amino aci
21	31	35.6	13	8	ADR99837	Adr99837 Citrullin
22	31	35.6	13	8	ADR99835	Adr99835 Human agg
23	30	34.5	10	2	AAR77850	Aar77850 [D-Ala 28
24	30	34.5	10	4	AAM10334	Aam10334 HLA-A1 de
25	30	34.5	10	4	AAM13196	Aam13196 HLA-A26 d
26	30	34.5	10	4	AAM09577	Aam09577 HLA-A *02
27	30	34.5	10	4	AAM10916	Aam10916 HLA-B *07
28	30	34.5	12	2	AAW85301	Aaw85301 Helper T-
29	30	34.5	12	5	AAU90801	Aau90801 Insulin/i
30	30	34.5	12	5	AAU88519	Aau88519 Insulin/i
31	30	34.5	12	5	AAU90701	Aau90701 Insulin/i
32	30	34.5	12	5	AAU88272	Aau88272 Insulin/i
33	30	34.5	12	6	ADA04537	Ada04537 IR/IGF-1R
34	30	34.5	12	6	ADA04920	Ada04920 Competito
35	30	34.5	12	7	ADH96133	Adh96133 Insulin r
36	30	34.5	12	7	ADH95750	Adh95750 Insulin/i
37	30	34.5	12	8	ADL68441	Adl68441 IGF-1R/IR
38	30	34.5	12	8	ADL68824	Adl68824 IGF-1R/IR
39	30	34.5	12	8	ADM38286	Adm38286 Represent
40	30	34.5	12	8	ADM38669	Adm38669 Insulin a
41	30	34.5	13	8	ADS92207	Ads92207 Chemosens
42	30	34.5	13	8	ADS92151	Ads92151 Chemosens
43	30	34.5	13	8	ADS92191	Ads92191 Chemosens
44	30	34.5	15	4	AAB86604	Aab86604 Human cyt
45	29	33.3	6	2	AAR29340	Aar29340 Endotheli

ALIGNMENTS

RESULT 1
AAW04853
ID AAW04853 standard; peptide; 15 AA.
XX
AC AAW04853;
XX
DT 27-AUG-2003 (revised)
DT 18-FEB-1997 (first entry)
XX
DE Internal fragment of influenza type A haemagglutinin protein.
XX
KW Tolerisation; self-epitope; antigen; autoimmune disease; autoantigen;
KW HLA; human leukocyte antigen; T-cell; thymocyte; pemphigus vulgaris;
KW desmoglein; multiple sclerosis; herpes simplex virus; adenovirus;
KW phosphomannomutase; human papillomavirus; Epstein-Barr virus;
KW DNA polymerase; influenza; haemagglutinin; reovirus; sigma protein.
XX
OS Influenza A virus.
XX
PN WO9627387-A1.
XX
PD 12-SEP-1996.
XX
PF 07-MAR-1996; 96WO-US003182.
XX
PR 07-MAR-1995; 95US-00400796.
XX
PA (HARD) HARVARD COLLEGE.
XX
PI Strominger JL, Wucherpfennig KW;
XX
DR WPI; 1996-425218/42.
XX
PT Pemphigus vulgaris auto-antigens and multiple sclerosis non-self antigens
PT - useful in disease treatment, and method for identification of other
PT self and non-self antigens implicated in auto-immune disease.
XX
PS Claim 2; Page 46; 58pp; English.
XX
CC Pharmaceutical preparations for tolerisation to antigens comprise either
CC an isolated human non-collagen or non-myslin basic protein (MBP)
CC polypeptide which is capable of tolerising an individual to an
CC autoantigen; or an isolated human pathogen polypeptide capable of
CC tolerising an individual to that polypeptide. In both cases, the
CC polypeptide (whether self or non-self) includes an amino acid sequence
CC corresponding to a sequence motif for a MHC class II protein, such as HLA
CC -DR, which is associated with a human autoimmune disease and which binds
CC to the polypeptide to activate autoreactive T-cells in individuals with

```
DE   Outer surface protein C (Fragment).
GN   Name=ospC;
OS   Borrelia garinii.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PBi;
RX   MEDLINE=97426044; PubMed=9282748;
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "he Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-373(1997).
DR   EMBL; U93699; AAC45533.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1019 MW;  4864C1A731A44333 CRC64;

  Query Match            25.3%;  Score 22;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 IKKNT 12
              :||||
Db          1 MKKNT 5


RESULT 14
Q9R3T0
ID   Q9R3T0      PRELIMINARY;      PRT;     9 AA.
AC   Q9R3T0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Outer surface protein C (Fragment).
GN   Name=ospC;
OS   Borrelia afzelii.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29518;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PKo, IP21, and J1;
RX   MEDLINE=97426044; PubMed=9282748;
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "he Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-373(1997).
DR   EMBL; U93698; AAC45531.1; -.
DR   EMBL; U93696; AAC45527.1; -.
DR   EMBL; U93697; AAC45529.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1005 MW;  4864C5B731A44333 CRC64;

  Query Match            25.3%;  Score 22;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 IKKNT 12
              :||||
Db          1 MKKNT 5


RESULT 15
Q9R792
ID   Q9R792      PRELIMINARY;      PRT;     9 AA.
AC   Q9R792;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Outer surface protein C (Fragment).
GN   Name=ospC;
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B31;
RX   MEDLINE=97426044; PubMed=9282748;
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "he Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-373(1997).
DR   EMBL; U93693; AAC45521.1; -.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1005 MW;  4864C5B731A44333 CRC64;

  Query Match            25.3%;  Score 22;  DB 2;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 IKKNT 12
              :||||
Db          1 MKKNT 5


Search completed: February 22, 2005, 09:38:03
Job time : 53.6667 secs
```

```
OC   Parapoxvirus.
OX   NCBI_TaxID=10258;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96010242; PubMed=7571439;
RA   Mercer A.A., Lyttle D.J., Whelan E.M., Fleming S.B., Sullivan J.T.;
RT   "The establishment of a genetic map of orf virus reveals a pattern of
RT   genomic organization that is highly conserved among divergent
RT   poxviruses.";
RL   Virology 212:698-704(1995).
DR   EMBL; U30336; AAA86400.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  1091 MW;  E24411A441F40446 CRC64;

  Query Match             27.6%;  Score 24;  DB 2;  Length 8;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          9 KKNTRY 14
              : ||||
Db          3 RHNTRY 8


RESULT 10
Q8SHM5
ID   Q8SHM5      PRELIMINARY;      PRT;    10 AA.
AC   Q8SHM5;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Bradypodion ventrale (Southern dwarf chameleon).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae;
OC   Bradypodion.
OX   NCBI_TaxID=179890;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22169767; PubMed=12182400; DOI=10.1006/mpev.2001.1076;
RA   Townsend T., Larson A.;
RT   "Molecular phylogenetics and mitochondrial genomic evolution in the
RT   chamaeleonidae (Reptilia, Squamata).";
RL   Mol. Phylogenet. Evol. 23:22-36(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Townsend T.M., Larson A.L.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF448732; AAL90478.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1353 MW;  DC218E2733640059 CRC64;

  Query Match             27.6%;  Score 24;  DB 2;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VWFIK 9
              :|||:
Db          1 MWFIR 5


RESULT 11
Q8SHM8
ID   Q8SHM8      PRELIMINARY;      PRT;    10 AA.
AC   Q8SHM8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Cytochrome c oxidase subunit I (Fragment).
GN   Name=COI;
OS   Bradypodion transvaalense (Transvaal dwarf chameleon).
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae;
OC   Bradypodion.
OX   NCBI_TaxID=179889;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22169767; PubMed=12182400; DOI=10.1006/mpev.2001.1076;
RA   Townsend T., Larson A.;
RT   "Molecular phylogenetics and mitochondrial genomic evolution in the
RT   chamaeleonidae (Reptilia, Squamata).";
RL   Mol. Phylogenet. Evol. 23:22-36(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Townsend T.M., Larson A.L.;
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF448731; AAL90475.1; -.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1353 MW;  DC218E2733640059 CRC64;

  Query Match             27.6%;  Score 24;  DB 2;  Length 10;
  Best Local Similarity   60.0%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 VWFIK 9
              :|||:
Db          1 MWFIR 5


RESULT 12
Q79B39
ID   Q79B39      PRELIMINARY;      PRT;     8 AA.
AC   Q79B39;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Outer surface protein C (Fragment).
GN   Name=ospC;
OS   Borrelia garinii.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=29519;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP90;
RX   MEDLINE=97426044; PubMed=9282748;
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "he Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-373(1997).
DR   EMBL; U93701; AAC45537.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  892 MW;  F4C5B731A4433336 CRC64;

  Query Match             25.3%;  Score 22;  DB 2;  Length 8;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+06;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          8 IKKNT 12
              :||||
Db          1 MKKNT 5


RESULT 13
O31363
ID   O31363      PRELIMINARY;      PRT;     9 AA.
AC   O31363;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21481970; PubMed=11598067;
RX   DOI=10.1128/IAI.69.11.6923-6930.2001;
RA   Wang L., Qu W., Reeves P.R.;
RT   "Sequence analysis of four Shigella boydii O-antigen loci: implication
RT   for Escherichia coli and Shigella relationships.";
RL   Infect. Immun. 69:6923-6930(2001).
DR   EMBL; AF409078; AAL27364.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1763 MW;  05FA72B5343C1234 CRC64;

  Query Match             31.0%;  Score 27;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          7 FIKKNTR 13
              ||:|| |
Db          3 FIRKNKR 9


RESULT 6
Q7SA62
ID   Q7SA62      PRELIMINARY;      PRT;    10 AA.
AC   Q7SA62;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Predicted protein.
GN   Name=NCU08327.1;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OR74A;
RA   Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA   Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA   Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA   Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA   Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA   Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA   Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA   Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA   Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA   Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA   DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA   Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA   Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA   Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT   "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL   Nature 0:0-0(2003).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABX01000206; EAA33304.1; -.
SQ   SEQUENCE   10 AA;  1284 MW;  D6D34CB72DC40059 CRC64;

  Query Match             29.9%;  Score 26;  DB 2;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          5 VWFIKKNTRY 14
              :|||:  : |
Db          1 MWFIRALSPY 10


RESULT 7
Q7M110
ID   Q7M110      PRELIMINARY;      PRT;    15 AA.
AC   Q7M110;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Endo-1,3-beta-glucanase (EC 3.2.1.-), 40k (Fragment).
OS   Bacillus circulans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=1397;
RN   [1]
RP   SEQUENCE.
RA   Fiske M.J., Tobey-Fincher K.L., Fuchs R.L.;
RT   "Cloning of two genes from Bacillus circulans WL-12 which encode 1,3-
RT   beta-glucanase activity.";
RL   J. Gen. Microbiol. 136:2377-2383(1990).
DR   PIR; B60763; B60763.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1740 MW;  2C6854FC5CACE1A6 CRC64;

  Query Match             28.7%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 NLVW 6
              ||||
Db          5 NLVW 8


RESULT 8
Q6LCW3
ID   Q6LCW3      PRELIMINARY;      PRT;    15 AA.
AC   Q6LCW3;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Transcription factor AP-2 isoform 2 (Fragment).
GN   Name=AP-2;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/4;
RX   MEDLINE=95269866; PubMed=7750631; DOI=10.1006/dbio.1995.1121;
RA   Meier P., Koedood M., Philipp J., Fontana A., Mitchell P.J.;
RT   "Alternative mRNAs encode multiple isoforms of transcription factor
RT   AP-2 during murine embryogenesis.";
RL   Dev. Biol. 169:1-14(1995).
DR   EMBL; U17289; AAA85677.1; -.
DR   GO; GO:0005667; C:transcription factor complex; TAS.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0003700; F:transcription factor activity; IDA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IDA.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1901 MW;  090AE08FEFC6136D CRC64;

  Query Match             28.7%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   27.3%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 LVWFIKKNTRY 14
              ::| :  | :|
Db          1 MLWKLTDNIKY 11


RESULT 9
Q84156
ID   Q84156      PRELIMINARY;      PRT;     8 AA.
AC   Q84156;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Similar to Vaccinia virus (Fragment).
OS   Orf virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
```

```
RC   STRAIN=LT2;
RX   MEDLINE=85088492; PubMed=6096856;
RA   Sacerdot C., Dessen P., Hershey J.W., Plumbridge J.A.,
RA   Grunberg-Manago M.;
RT   "Sequence of the initiation factor IF2 gene: unusual protein features
RT   and homologies with elongation factors.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:7787-7791(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=94156845; PubMed=8113180;
RA   Craven M.G., Granston A.E., Schauer A.T., Zheng C., Gray T.A.,
RA   Friedman D.I.;
RT   "Escherichia coli-Salmonella typhimurium hybrid nusA genes:
RT   identification of a short motif required for action of the lambda N
RT   transcription antitermination protein.";
RL   J. Bacteriol. 176:1394-1404(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE=95212911; PubMed=7535280;
RA   Dammel C.S., Noller H.F.;
RT   "Suppression of a cold-sensitive mutation in 16S rRNA by
RT   overexpression of a novel ribosome binding factor, RbfA.";
RL   Genes Dev. 9:626-637(1995).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RA   Hedegaard J., Steffensen S.A., Norskow-Lauritsen N., Mortensen K.K.,
RA   Sperling-Petersen H.U.;
RT   "Identification of Enterobacteriaceae by partial sequencing of the
RT   gene encoding translation initiation factor 2.";
RL   Int. J. Syst. Bacteriol. 48:1531-1538(1999).
DR   EMBL; AJ002552; CAA05548.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   15 AA;  1710 MW;  1C6AD0FD23B00DCC CRC64;

  Query Match             33.3%;  Score 29;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 5.9e+02;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          2 RNLVWF 7
              ||: ||
Db          6 RNICWF 11


RESULT 3
APE_CAPGI
ID   APE_CAPGI      STANDARD;      PRT;    10 AA.
AC   P80474;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Aminopeptidase (EC 3.4.11.-) (Fragment).
OS   Capnocytophaga gingivalis.
OC   Bacteria; Bacteroidetes; Flavobacteria; Flavobacteriales;
OC   Flavobacteriaceae; Capnocytophaga.
OX   NCBI_TaxID=1017;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=ATCC 33624;
RX   MEDLINE=96118234; PubMed=8574402;
RA   Spratt D.A., Greenman J., Schaffer A.G.;
RT   "Capnocytophaga gingivalis aminopeptidase: a potential virulence
RT   factor.";
RL   Microbiology 141:3087-3093(1995).
CC   -!- FUNCTION: Aminopeptidase which hydrolyzes substrates with free N-
CC       terminal amino acid residues but not N-terminal blocked ones.
CC       Optimum activity is measured at pH 7.5. May be important in the
CC       nutrition and pathogenesis of the organism in the human oral
CC       cavity.
CC   -!- COFACTOR: Requires magnesium or calcium.
KW   Aminopeptidase; Calcium; Direct protein sequencing; Hydrolase;
KW   Magnesium.
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1306 MW;  00C0A6DB43772694 CRC64;

  Query Match             31.0%;  Score 27;  DB 1;  Length 10;
  Best Local Similarity   25.0%;  Pred. No. 8.5e+02;
  Matches    2;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 NLVWFIKK 10
              |::|:: :
Db          3 NMLWYVXR 10


RESULT 4
Q7RQJ4
ID   Q7RQJ4      PRELIMINARY;      PRT;    12 AA.
AC   Q7RQJ4;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   Name=PY01103;
OS   Plasmodium yoelii yoelii.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=73239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=17XNL;
RX   PubMed=12368865; DOI=10.1038/nature01099;
RA   Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA   Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA   Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA   Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA   Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA   Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA   Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA   van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA   Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA   Carucci D.J.;
RT   "Genome sequence and comparative analysis of the model rodent malaria
RT   parasite Plasmodium yoelii yoelii.";
RL   Nature 419:512-519(2002).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABL01000291; EAA20281.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   12 AA;  1409 MW;  96C174ED7A3059D1 CRC64;

  Query Match             31.0%;  Score 27;  DB 2;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 NLVWFIKKN 11
              ||  ||| |
Db          2 NLSTFIKNN 10


RESULT 5
Q93CI1
ID   Q93CI1      PRELIMINARY;      PRT;    14 AA.
AC   Q93CI1;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   O-antigen polymerase (Fragment).
GN   Name=wzy;
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 08:53:13 ; Search time 52.6 Seconds
(without alignments)
146.030 Million cell updates/sec

Title: US-08-991-628-13
Perfect score: 87
Sequence: 1 *YRNLVWFIKKNTRYP* 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 6622

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	33.3	13	2	O99188	O99188 rivulus cau
2	29	33.3	15	2	Q9ZF32	Q9zf32 salmonella
3	27	31.0	10	1	APE_CAPGI	P80474 capnocytoph
4	27	31.0	12	2	Q7RQJ4	Q7rqj4 plasmodium
5	27	31.0	14	2	Q93CI1	Q93ci1 escherichia
6	26	29.9	10	2	Q7SA62	Q7sa62 neurospora
7	25	28.7	15	2	Q7M110	Q7m110 bacillus ci
8	25	28.7	15	2	Q6LCW3	Q6lcw3 mus musculu
9	24	27.6	8	2	Q84156	Q84156 orf virus.
10	24	27.6	10	2	Q8SHM5	Q8shm5 bradypodion
11	24	27.6	10	2	Q8SHM8	Q8shm8 bradypodion
12	22	25.3	8	2	Q79B39	Q79b39 borrelia ga
13	22	25.3	9	2	O31363	O31363 borrelia ga
14	22	25.3	9	2	Q9R3T0	Q9r3t0 borrelia af
15	22	25.3	9	2	Q9R792	Q9r792 borrelia bu
16	22	25.3	10	2	P83067	P83067 bacillus ce
17	22	25.3	10	2	Q9R791	Q9r791 borrelia af
18	22	25.3	11	2	Q9R790	Q9r790 borrelia ga
19	22	25.3	13	2	Q64FK2	Q64fk2 homo sapien
20	22	25.3	14	2	Q7RR89	Q7rr89 plasmodium
21	22	25.3	14	2	Q9QW75	Q9qw75 mus sp. hom
22	22	25.3	15	1	UC16_MAIZE	P80622 zea mays (m
23	22	25.3	15	2	O35411	O35411 mus musculu
24	21	24.1	10	2	Q79AV7	Q79av7 klebsiella
25	21	24.1	11	2	Q9AIZ7	Q9aiz7 carsonella
26	21	24.1	13	2	Q6LDS1	Q6lds1 bacteriopha
27	21	24.1	15	1	UC08_MAIZE	P80614 zea mays (m
28	21	24.1	15	2	Q9UWM1	Q9uwm1 methanospir
29	21	24.1	15	2	Q35795	Q35795 saccharomyc
30	20.5	23.6	14	2	Q89818	Q89818 murine minu
31	20	23.0	7	2	Q8GL00	Q8gl00 borrelia bu
32	20	23.0	7	2	Q8GL04	Q8gl04 borrelia bu
33	20	23.0	8	2	Q95M23	Q95m23 sus scrofa
34	20	23.0	8	2	Q8G940	Q8g940 borrelia bu
35	20	23.0	9	2	Q28121	Q28121 bos taurus
36	20	23.0	9	2	Q8GL26	Q8gl26 borrelia bu
37	20	23.0	9	2	Q9R9C4	Q9r9c4 borrelia bu
38	20	23.0	10	2	Q8SHA2	Q8sha2 brookesia b
39	20	23.0	10	2	P82438	P82438 nicotiana t
40	20	23.0	10	2	Q8G8W5	Q8g8w5 borrelia bu
41	20	23.0	11	2	Q25916	Q25916 plasmodium
42	20	23.0	11	2	Q8GL24	Q8gl24 borrelia bu
43	20	23.0	12	2	Q8IVH0	Q8ivh0 homo sapien
44	20	23.0	12	2	Q7M1H0	Q7m1h0 leonurus ar
45	20	23.0	13	1	TEML_RANTE	P57104 rana tempor

ALIGNMENTS

RESULT 1
O99188
ID O99188 PRELIMINARY; PRT; 13 AA.
AC O99188;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cytochrome oxidase I (Fragment).
GN Name=COI;
OS Rivulus caudomarginatus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Aplocheilidae; Rivulinae; Rivulus.
OX NCBI_TaxID=60318;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20072928; PubMed=10603257; DOI=10.1006/mpev.1999.0656;
RA Murphy W.J., Thomerson J.E., Collier G.E.;
RT "Phylogeny of the Neotropical killifish family Rivulidae
RT (Cyprinodontiformes, Aplocheiloidei) inferred from mitochondrial DNA
RT sequences.";
RL Mol. Phylogenet. Evol. 13:289-301(1999).
DR EMBL; AF002597; AAD01080.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 13 AA; 1705 MW; 404DF35AEFFE79C7 CRC64;

Query Match 33.3%; Score 29; DB 2; Length 13;
Best Local Similarity 57.1%; Pred. No. 5.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 YRNLVWF 7
              |::| ||
Db          3 YQHLFWF 9
```

RESULT 2
Q9ZF32
ID Q9ZF32 PRELIMINARY; PRT; 15 AA.
AC Q9ZF32;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NusA protein (Fragment).
GN Name=nusA;
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.

```
A;Accession: E60396
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-13 <LIM>
A;Cross-references: GB:M31305

  Query Match             23.0%;  Score 20;  DB 2;  Length 13;
  Best Local Similarity   75.0%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 YRNL 4
              |||:
Db         10 YRNI 13


RESULT 15
PN0662
dystrophin-associated glycoprotein A3a-I - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 07-May-1999
C;Accession: PN0662
R;Yoshida, M.; Mizuno, Y.; Nonaka, I.; Ozawa, E.
J. Biochem. 114, 634-639, 1993
A;Title: A dystrophin-associated glycoprotein, A3a (one of 43DAG doublets), is retained
A;Reference number: PN0662; MUID:94156881; PMID:8113213
A;Accession: PN0662
A;Molecule type: protein
A;Residues: 1-15 <YOS>
C;Comment: This protein is retained in Duchenne type muscular dystrophy muscle.
C;Keywords: glycoprotein; skeletal muscle

  Query Match             23.0%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          7 FIKK 10
              ||||
Db          5 FIKK 8


Search completed: February 22, 2005, 09:46:29
Job time : 11.1333 secs
```

```
  Best Local Similarity   75.0%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 TRYP 15
              |:||
Db          6 TKYP 9


RESULT 9
A48372
benzoyl-CoA ligase - Methanospirillum hungatei (fragment)
C;Species: Methanospirillum hungatei
C;Date: 19-Nov-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A48372
R;Auburger, G.; Winter, J.
Appl. Microbiol. Biotechnol. 37, 789-795, 1992
A;Title: Purification and characterization of benzoyl-CoA ligase from a syntrophic, benz
A;Reference number: A48372; MUID:93040109; PMID:1369492
A;Accession: A48372
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <AUB>
A;Cross-references: UNIPROT:Q9UWM1
A;Note: sequence extracted from NCBI backbone (NCBIP:118357)

  Query Match             24.1%;  Score 21;  DB 2;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 3.2e+03;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 YRNLVWF 7
              | :  ||
Db          9 YNSADWF 15


RESULT 10
I46023
growth hormone receptor - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I46023
R;Hauser, S.D.; McGrath, M.F.; Collier, R.J.; Krivi, G.G.
Mol. Cell. Endocrinol. 72, 187-200, 1990
A;Title: Cloning and in vivo expression of bovine growth hormone receptor mRNA.
A;Reference number: I46023; MUID:91146804; PMID:2289631
A;Accession: I46023
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-9 <HAU>
A;Cross-references: UNIPROT:Q28121; EMBL:U24113; NID:g775221; PIDN:AAA91014.1; PID:g7752
C;Genetics:
A;Gene: GHR

  Query Match             23.0%;  Score 20;  DB 2;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 2.8e+05;
  Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          6 WFI 8
              ||:
Db          5 WFV 7


RESULT 11
B44818
extracellular lipase - Pseudomonas aeruginosa (fragment)
C;Species: Pseudomonas aeruginosa
C;Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: B44818
R;Gilbert, E.J.; Cornish, A.; Jones, C.W.
J. Gen. Microbiol. 137, 2223-2229, 1991
A;Title: Purification and properties of extracellular lipase from Pseudomonas aeruginosa
A;Reference number: A44818; MUID:92085040; PMID:1748875
A;Accession: B44818
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <GIL>
A;Experimental source: strain EF228
A;Note: sequence extracted from NCBI backbone (NCBIP:70393)

  Query Match             23.0%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 3.8e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 TRYP 15
              |:||
Db          5 TQYP 8


RESULT 12
S69123
proton-translocating transhydrogenase - Rhodospirillum rubrum (fragment)
C;Species: Rhodospirillum rubrum
C;Date: 10-Mar-1998 #sequence_revision 24-Apr-1998 #text_change 24-Apr-1998
C;Accession: S69123
R;Diggle, C.; Hutton, M.; Jones, G.R.; Thomas, C.M.; Jackson, J.B.
Eur. J. Biochem. 228, 719-726, 1995
A;Title: Properties of the soluble polypeptide of the proton-translocating transhydrogen
A;Reference number: S69123; MUID:95255277; PMID:7737169
A;Accession: S69123
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <DIG>

  Query Match             23.0%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   25.0%;  Pred. No. 3.8e+03;
  Matches    2;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 NLVWFIKK 10
              ::|| :::
Db          3 DVVWKVQR 10


RESULT 13
JU0356
cycloleonurinin - sagebrush motherwort
C;Species: Leonurus artemisia (sagebrush motherwort)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
C;Accession: JU0356
R;Kinoshita, K.; Tanaka, J.; Kuroda, K.; Koyama, K.; Natori, S.; Kinoshita, T.
Chem. Pharm. Bull. 39, 712-715, 1991
A;Title: Cycloleonurinin, a cyclic peptide from Leonuri fructus.
A;Reference number: JU0356; MUID:91300597; PMID:2070452
A;Accession: JU0356
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <KIN>
A;Cross-references: UNIPROT:Q7M1H0

  Query Match             23.0%;  Score 20;  DB 2;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 3.8e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         12 TRYP 15
              |:||
Db          7 TQYP 10


RESULT 14
E60396
antigen 7H8/2 - malaria parasite (Plasmodium falciparum) (fragments)
C;Species: Plasmodium falciparum
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jun-2000
C;Accession: E60396
R;Limpaiboon, T.; Taylor, D.W.; Jones, G.; Geysen, H.M.; Saul, A.
Southeast Asian J. Trop. Med. Public Health 21, 388-396, 1990
A;Title: Characterization of a Plasmodium falciparum epitope recognized by a monoclonal
A;Reference number: A60396; MUID:91164876; PMID:1706114
```

PN0649
pullulanase (EC 3.2.1.41) - Bacillus sp. (strain S-1) (fragment)
C;Species: Bacillus sp.
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 03-Jun-2002
C;Accession: PN0649
R;Kim, C.H.; Choi, H.I.; Lee, D.S.
Biosci. Biotechnol. Biochem. 57, 1632-1637, 1993
A;Title: Purification and biochemical properties of an alkaline pullulanase from alkalop
A;Reference number: PN0649; MUID:94080025; PMID:7764261
A;Accession: PN0649
A;Molecule type: protein
A;Residues: 1-7 <KIM>
C;Comment: This enzyme is used togther with glucoamylase to improve the efficiency of st
ntent in high maltose syrups.
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

  Query Match 27.6%; Score 24; DB 2; Length 7;
  Best Local Similarity 42.9%; Pred. No. 2.8e+05;
  Matches 3; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 YRNLVWF 7
  : |: ||
Db 1 FLNMSWF 7

RESULT 4
C48401
ribosomal protein L1 - Thermus aquaticus (fragment)
C;Species: Thermus aquaticus
C;Date: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: C48401
R;Garber, M.B.; Agalarov, S.C.; Eliseikina, I.A.; Fomenkova, N.P.; Nikonov, S.V.; Sedeln
Biochimie 74, 327-336, 1992
A;Title: Ribosomal proteins from Thermus thermophilus for structural investigations.
A;Reference number: A48401; MUID:92345325; PMID:1637860
A;Accession: C48401
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <GAR>
A;Note: sequence extracted from NCBI backbone (NCBIP:109932)

  Query Match 27.6%; Score 24; DB 2; Length 15;
  Best Local Similarity 57.1%; Pred. No. 1e+03;
  Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 8 IKKNTRY 14
  :| | ||
Db 1 VKTNKRY 7

RESULT 5
PC1215
homeotic protein EgHbx5 - tapeworm (Echinococcus granulosus) (fragment)
C;Species: Echinococcus granulosus
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 30-Sep-1993
C;Accession: PC1215
R;Oliver, G.; Vispo, M.; Mailhos, A.; Martinez, C.; Sosa-Pineda, B.; Fielitz, W.; Ehrlic
Gene 121, 337-342, 1992
A;Title: Homeoboxes in flatworms.
A;Reference number: JC1386; MUID:93077050; PMID:1359988
A;Accession: PC1215
A;Molecule type: DNA
A;Residues: 1-14 <OLI>
A;Cross-references: EMBL:X66821
C;Keywords: homeobox

  Query Match 26.4%; Score 23; DB 2; Length 14;
  Best Local Similarity 33.3%; Pred. No. 1.4e+03;
  Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 5 VWFIKKNTR 13
  ||| : ::
Db 1 VWFQNRRSK 9

RESULT 6
PH1626
Ig H chain V-D-J region (clone B-less 118) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1626
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A;Reference number: PH1580; MUID:93301609; PMID:8315387
A;Accession: PH1626
A;Molecule type: DNA
A;Residues: 1-14 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Keywords: immunoglobulin

  Query Match 25.3%; Score 22; DB 2; Length 14;
  Best Local Similarity 50.0%; Pred. No. 2.1e+03;
  Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YRNLVW 6
  | :|:|
Db 8 YGDLLW 13

RESULT 7
A47421
leukotriene B-4 12-hydroxydehydrogenase (EC 1.1.1.-) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 09-Jul-2004
C;Accession: A47421
R;Yokomizo, T.; Izumi, T.; Takahashi, T.; Kasama, T.; Kobayashi, Y.; Sato, F.; Taketani,
J. Biol. Chem. 268, 18128-18135, 1993
A;Title: Enzymatic inactivation of leukotriene B-4 by a novel enzyme found in the porcin
A;Reference number: A47421; MUID:93352633; PMID:8394361
A;Accession: A47421
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <YOK>
A;Cross-references: UNIPROT:O62642
C;Keywords: oxidoreductase

  Query Match 25.3%; Score 22; DB 2; Length 14;
  Best Local Similarity 50.0%; Pred. No. 2.1e+03;
  Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 6 WFIKKN 11
  | :||:
Db 6 WTLKKH 11

RESULT 8
A44818
extracellular lipase - Pseudomonas aeruginosa (fragment)
C;Species: Pseudomonas aeruginosa
C;Date: 31-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A44818
R;Gilbert, E.J.; Cornish, A.; Jones, C.W.
J. Gen. Microbiol. 137, 2223-2229, 1991
A;Title: Purification and properties of extracellular lipase from Pseudomonas aeruginosa
A;Reference number: A44818; MUID:92085040; PMID:1748875
A;Accession: A44818
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-13 <GIL>
A;Cross-references: UNIPROT:Q9L6C7; UNIPROT:P95419
A;Experimental source: strain TE3285
A;Note: sequence extracted from NCBI backbone (NCBIP:70395)
C;Superfamily: Pseudomonas triacylglycerol lipase

  Query Match 24.1%; Score 21; DB 2; Length 13;

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: February 22, 2005, 09:08:09 ; Search time 11.1333 Seconds
(without alignments)
129.633 Million cell updates/sec

Title: US-08-991-628-13
Perfect score: 87
Sequence: 1 YRNLVWFIKKNTRYP 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 2523

Minimum DB seq length: 0
Maximum DB seq length: 15

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	28.7	10	2	PT0310	Ig heavy chain CRD
2	25	28.7	15	2	B60763	endo-1,3-beta-gluc
3	24	27.6	7	2	PN0649	pullulanase (EC 3.
4	24	27.6	15	2	C48401	ribosomal protein
5	23	26.4	14	2	PC1215	homeotic protein E
6	22	25.3	14	2	PH1626	Ig H chain V-D-J r
7	22	25.3	14	2	A47421	leukotriene B-4 12
8	21	24.1	13	2	A44818	extracellular lipa
9	21	24.1	15	2	A48372	benzoyl-CoA ligase
10	20	23.0	9	2	I46023	growth hormone rec
11	20	23.0	12	2	B44818	extracellular lipa
12	20	23.0	12	2	S69123	proton-translocati
13	20	23.0	12	2	JU0356	cycloleonurinin -
14	20	23.0	13	2	E60396	antigen 7H8/2 - ma
15	20	23.0	15	2	PN0662	dystrophin-associa
16	19	21.8	7	2	PX0008	glucuronosyltransf
17	19	21.8	10	2	F41839	ribosomal protein
18	19	21.8	11	2	T12264	cytochrome-c oxida
19	19	21.8	11	2	T12253	cytochrome-c oxida
20	19	21.8	11	2	T17081	cytochrome-c oxida
21	19	21.8	12	2	S25056	Ig heavy chain - m
22	19	21.8	15	2	PH1329	Ig heavy chain DJ
23	19	21.8	15	2	PQ0074	T-cell receptor be
24	19	21.8	15	2	B35389	urease (EC 3.5.1.5
25	18	20.7	8	2	S19288	acylase - Kluyvera
26	18	20.7	8	2	S43971	tumor-associated a
27	18	20.7	11	2	I54193	Rhesus blood group
28	18	20.7	12	2	I64829	gene HEXA protein
29	18	20.7	13	2	S47358	T-cell antigen rec
30	18	20.7	13	2	PC2369	unidentified 85K p
31	18	20.7	14	2	A60840	somatostatin I - E
32	18	20.7	14	2	C60414	somatostatin - sli
33	18	20.7	14	2	B60842	somatostatin I - c
34	18	20.7	14	2	S00172	somatostatin I - s
35	18	20.7	14	2	I54284	C1-inhibitor - hum
36	18	20.7	14	2	I64815	carbonic anhydrase
37	18	20.7	14	2	PH1757	T cell receptor al
38	18	20.7	14	2	PH1758	T cell receptor al
39	18	20.7	14	2	PH1759	T cell receptor al
40	18	20.7	14	2	PH1766	T cell receptor al
41	18	20.7	14	2	PH1767	T cell receptor al
42	18	20.7	14	2	PH1768	T cell receptor al
43	18	20.7	14	2	PH1769	T cell receptor al
44	18	20.7	14	2	PH1628	Ig H chain V-D-J r
45	18	20.7	14	2	A44920	2-halobenzoate 1,2

ALIGNMENTS

RESULT 1
PT0310
Ig heavy chain CRD3 region (clone 6-97) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0310
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0310
A;Molecule type: DNA
A;Residues: 1-10 <YAM>
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 28.7%; Score 25; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVWF 7
||||
Db 3 LVWF 6

RESULT 2
B60763
endo-1,3-beta-glucanase (EC 3.2.1.-), 40k - Bacillus circulans (strain WL-12) (fragment)
C;Species: Bacillus circulans
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 09-Jul-2004
C;Accession: B60763
R;Fiske, M.J.; Tobey-Fincher, K.L.; Fuchs, R.L.
J. Gen. Microbiol. 136, 2377-2383, 1990
A;Title: Cloning of two genes from Bacillus circulans WL-12 which encode 1,3-beta-glucai
A;Reference number: A60763; MUID:91178514; PMID:2127800
A;Accession: B60763
A;Molecule type: protein
A;Residues: 1-15 <FIS>
A;Cross-references: UNIPROT:Q7M110
C;Comment: This bacillus produces up to six different 1,3-beta-glucanases for growth on
C;Keywords: glycosidase; hydrolase

Query Match 28.7%; Score 25; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 NLVW 6
||||
Db 5 NLVW 8

RESULT 3

```
CC   and methods for the treatment and diagnosis of chlamydial infection. The
CC   compounds provided include polypeptides and fusion proteins comprising
CC   immunogenic portions of Chlamydia antigens and DNA sequences encoding
CC   such polypeptides. They are useful for vaccinating against chlamydial
CC   infection, which causes pelvic inflammatory disease, trachoma, acute
CC   respiratory tract infections, atherosclerosis and heart disease
XX
SQ   Sequence 15 AA;

  Query Match             35.3%;  Score 30;  DB 4;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 VYHFVKKH 11
              |: ::|||
Db          2 VWEYIKKH 9


RESULT 15
ABB94132
ID   ABB94132 standard; peptide; 15 AA.
XX
AC   ABB94132;
XX
DT   05-JUN-2002  (first entry)
XX
DE   Chlamydia peptide sequence SEQ ID NO:103.
XX
KW   Chlamydial infection; Chlamydia; vaccine; detection; diagnosis; antigen;
KW   antibacterial; immunostimulant; immune response;
KW   Chlamydia-specific T-cell response.
XX
OS   Chlamydia sp.
OS   Synthetic.
XX
PN   WO200208267-A2.
XX
PD   31-JAN-2002.
XX
PF   20-JUL-2001; 2001WO-US023121.
XX
PR   20-JUL-2000; 2000US-00620412.
PR   23-APR-2001; 2001US-00841132.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Fling SP,  Skeiky YAW,  Probst P,  Bhatia A;
XX
DR   WPI; 2002-179901/23.
XX
PT   Novel compositions comprising Chlamydia Cap1 protein and its use in the
PT   treatment of Chlamydia infection.
XX
PS   Disclosure; Page 191; 537pp; English.
XX
CC   The present invention describes compositions comprising a Chlamydia Cap1
CC   protein and methods for the diagnosis and therapy of Chlamydia infection.
CC   Chlamydia DNA and protein sequences from the present invention can have
CC   antibacterial and immunostimulant activities, and can be used in
CC   vaccines. Compounds from the present invention can be used for eliciting
CC   an immune response, specifically stimulating a Chlamydia-specific T-cell
CC   response or inhibiting the development of a Chlamydia infection in an
CC   animal. Methods from the present invention can be used: for detecting the
CC   presence of Chlamydia in a patient; to stimulate and/or expand T cells
CC   specific for a Chlamydia protein; and for treatment of a Chlamydia
CC   infection. ABL92394 to ABL92709 and ABB94096 to ABB94374 represent
CC   sequences used in the exemplification of the present invention
XX
SQ   Sequence 15 AA;

  Query Match             35.3%;  Score 30;  DB 5;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 VYHFVKKH 11
              |: ::|||
Db          2 VWEYIKKH 9


Search completed: February 22, 2005, 09:24:52
Job time : 67.6667 secs
```

XX
PR 19-JUN-2002; 2002DE-01027238.
XX
PA (WELA) WELLA AG.
XX
PI Rothe H, Aeby P, Goettel O, Braun H;
XX
DR WPI; 2004-108357/11.
XX
PT Cosmetic composition, especially hair dye, comprises a peptide with high
PT affinity for keratinic materials covalently linked to a cosmetic active
PT agent.
XX
PS Disclosure; Page 20; 30pp; German.
XX
CC This invention describes a novel cosmetic composition which comprises a
CC peptide with high affinity for keratinic materials covalently linked to a
CC cosmetic active agent (including dyes, hair conditioners, combability
CC improvers, surfactants, amidoamines, betaine esters, esterquats, fatty
CC alcohols, moisturisers, vitamins, provitamins, bounce improvers,
CC betaines, sugars and UV filters. Preferred agents include polymers,
CC silicone polyols and chitosan. The composition is useful for cosmetic
CC treatment of human or animal keratinic materials, especially as a hair
CC dye. The peptide targets the active agent to keratinic materials. The
CC invention describes a dye/amino acid conjugate prepared by reacting
CC Reactive Red 2 with DL-lysine hydrochloride. ADI32195-ADI32295 represent
CC peptide linkers used to make the cosmetic constructs described in the
CC invention.
XX
SQ Sequence 13 AA;

Query Match 35.3%; Score 30; DB 8; Length 13;
Best Local Similarity 50.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 6 HFVKKHVHES 15
| ||| |::
Db 2 HHKKKHHHKT 11

RESULT 13
AAB13485
ID AAB13485 standard; peptide; 15 AA.
XX
AC AAB13485;
XX
DT 02-FEB-2001 (first entry)
XX
DE C. trachomatis Ct-Swib36-50 peptide.
XX
KW Chlamydial infection; sexually transmitted disease;
KW pelvic inflammatory disease; PID; tubal obstruction; infertility;
KW trachoma; blindness; acute respiratory tract infection; atherosclerosis;
KW coronary heart disease; antibacterial.
XX
OS Chlamydia trachomatis.
XX
PN WO200034483-A2.
XX
PD 15-JUN-2000.
XX
PF 08-DEC-1999; 99WO-US029012.
XX
PR 08-DEC-1998; 98US-00208277.
PR 08-APR-1999; 99US-00288594.
PR 01-OCT-1999; 99US-00410568.
PR 22-OCT-1999; 99US-00426571.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EJ;
XX
DR WPI; 2000-431303/37.
XX
PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection
PT comprises immunogenic portion of Chlamydia antigen, which comprises amino
PT acid sequence encoded by polynucleotide sequence.
XX
PS Claim 2; Page 148; 256pp; English.
XX
CC The present invention relates to new nucleic acid sequences and the
CC proteins encoded by the nucleic acid sequences. The encoded proteins
CC comprise an immunogenic portion of a Chlamydia antigen. The encoded
CC proteins are useful for the serodiagnosis and treatment of Chlamydia
CC infection. Chlamdiae are intracellular bacterial pathogens that are
CC responsible for a wide variety of human infections. C. trachomatis
CC infection is one of the most common sexually transmitted diseases and can
CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction
CC and infertility. Trachoma due to ocular infection with C. trachomatis is
CC the leading cause of preventable blindness worldwide. C. pneumonia is a
CC major cause of acute respiratory tract infections in humans and is also
CC thought to play a role in the pathogenesis of atherosclerosis and
CC coronary heart disease. The present sequence is a protein isolated in the
CC present invention
XX
SQ Sequence 15 AA;

Query Match 35.3%; Score 30; DB 3; Length 15;
Best Local Similarity 50.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VYHFVKKH 11
|: ::|||
Db 2 VWEYIKKH 9

RESULT 14
AAG83161
ID AAG83161 standard; peptide; 15 AA.
XX
AC AAG83161;
XX
DT 05-SEP-2001 (first entry)
XX
DE Chlamydia trachomatis Ct-Swib 36-50 peptide.
XX
KW Chlamydia; vaccine; infection; fusion protein; antigen;
KW pelvic inflammatory disease; trachoma; atherosclerosis; heart disease;
KW acute respiratory tract infection; Cap1; CT529; OMCB;
KW polymorphic membrane protein; pmp; thiol specific antioxidant; TSA.
XX
OS Chlamydia trachomatis.
XX
PN WO200140474-A2.
XX
PD 07-JUN-2001.
XX
PF 04-DEC-2000; 2000WO-US032919.
XX
PR 03-DEC-1999; 99US-00454684.
PR 19-APR-2000; 2000US-00556877.
PR 20-JUN-2000; 2000US-00598419.
XX
PA (CORI-) CORIXA CORP.
XX
PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Scholler J;
XX
DR WPI; 2001-374831/39.
XX
PT Chlamydia polypeptides and fusion proteins useful for preventing pelvic
PT inflammatory disease, trachoma, acute respiratory tract infections,
PT atherosclerosis and heart disease.
XX
PS Disclosure; Page 157; 295pp; English.
XX
CC The present peptide is provided in a specification relating to compounds

CC from the F2a and F2c regions of the proteins GDNF, neurturin and artemin, CC from humans, mice or rats. This type of protein activates the growth CC factor receptor alpha1-Ret protein-tyrosine kinase (GFRalpha1-RET), but CC does not substantially activate GFRalpha2-RET or GFRalpha3-RET. The CC growth factors and nucleic acids encoding them are useful for providing CC trophic support to a mammalian cell and/or for producing differentiation CC of a mammalian cell, in a patient suffering from peripheral neuropathy, CC amyotrophic lateral sclerosis, Alzheimer's disease, Parkinson's disease, CC Huntington's disease, diabetes, acquired immunodeficiency syndrome CC (AIDS), ischaemic stroke, acute brain injury, acute spinal cord injury, CC multiple sclerosis, nervous system tumours (e.g. neuroblastomas), or CC enteric diseases such as idiopathic constipation. The sequences are also CC useful for preventing or treating cellular degeneration or insufficiency CC in an individual, suffering from eosinopaenia, basopaenia, lymphopaenia, CC monocytopaenia, neutropaenia, anaemia, thrombocytopaenia, cardiac muscle CC degeneration, or congestive heart failure. The growth factors are also CC useful for promoting the survival of peripheral and central neuronal CC populations in vivo or in vitro
XX
SQ Sequence 7 AA;

Query Match 36.5%; Score 31; DB 4; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.8e+06;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy          5 YHFVKKH 11
              || ::||
Db          1 YHILRKH 7
```

RESULT 9
ADB81184
ID ADB81184 standard; peptide; 13 AA.
XX
AC ADB81184;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human pancreatic/heart tissue quantified by LC-MS analysis, SEQ ID No 75.
XX
KW quantitating relative abundance; liquid phase tandem mass spectroscopy;
KW data set; peptide profile; toxicology analysis; classification; human.
XX
OS Homo.
OS sapiens.
XX
PN WO200297703-A2.
XX
PD 05-DEC-2002.
XX
PF 30-MAY-2002; 2002WO-CA000801.
XX
PR 30-MAY-2001; 2001CA-02349265.
XX
PA (EMIL/) EMILI A.
PA (CAGN/) CAGNEY G.
XX
PI Emili A, Cagney G;
XX
DR WPI; 2003-175129/17.
XX
PT Identifying and quantifying proteins/peptide sequences, comprises
PT identifying peptides by liquid phase tandem mass spectroscopy sequencing
PT and compiling data set/peptide profile containing collection of peptide
PT sequences.
XX
PS Example 3; Page 49; 84pp; English.
XX
CC The invention relates to a novel method for identifying constituent
CC proteins/peptide sequences or quantitating relative abundance of proteins
CC in two samples, for a cell type, tissue or pathological sample using a
CC database having peptide profile libraries with multiple protein/peptide
CC sequences. The method comprises identifying peptide species by liquid
CC phase tandem mass spectroscopy sequencing and compiling a data
CC set/peptide profile. The novel method is useful in toxicology analysis.
CC In the method of the invention, there is no requirement for prior
CC knowledge about the functions of the responsive peptides or parental
CC proteins. Protein functions deduced from comparisons of profiles in a
CC database can be derived from very subtle physiological responses. For
CC instance, even though peptide levels may change only slightly in response
CC to an experimental treatment, coordinate changes among many measured
CC peptide abundances can be sufficient to characterise that phenotype. The
CC large number of peptides measured make it unlikely that an unrelated
CC physiological state will have an identical profile, even though this may
CC not be apparent when using conventional experiments that measure the
CC levels of one or a few proteins. Closely related profiles can be classed
CC together, thus improving the understanding of the underlying biological
CC basis of the classifications. This sequence represents a peptide
CC sequenced and quantified from a human lung tissue lysate sample in a
CC single liquid chromatography-mass spectrometry (LC-MS) analysis, used in
CC the method for the measurement of protein relative abundance in complex
CC mixtures.
XX
SQ Sequence 13 AA;

Query Match 36.5%; Score 31; DB 7; Length 13;
Best Local Similarity 50.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 TGGVYHFVKK 10
              |||: |  |:
Db          4 TGGIAHLFKQ 13
```

RESULT 10
AAW11603
ID AAW11603 standard; peptide; 11 AA.
XX
AC AAW11603;
XX
DT 22-SEP-1997 (first entry)
XX
DE Linear ZnF peptide.
XX
KW Conformationally-restricted polypeptide; alpha-helical coiled coil;
KW combinatorial library; macromolecular ligand; peptidomimetic; antibody;
KW ion channel.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "In acetyl form"
FT Modified-site 11
FT /note= "In amide form"
XX
PN WO9700267-A1.
XX
PD 03-JAN-1997.
XX
PF 14-JUN-1996; 96WO-CA000403.
XX
PR 16-JUN-1995; 95US-00491527.
XX
PA (PENC-) PENCE INC.
XX
PI Houston ME, Hodges RS;
XX
DR WPI; 1997-077485/07.
XX
PT Combinatorial library of conformationally-restricted polypeptide(s) -
PT used to screen for cpds. that react specifically with a macromolecular
PT ligand and for developing peptidomimetics.
XX
PS Example 6; Fig 9; 66pp; English.
XX

```
CC   The present sequence represents a linear ZnF peptide. This peptide was
CC   used to show that peptides with lactam bridges inhibit antibody binding
CC   more effectively than peptides without lactam bridges. The present
CC   peptide could inhibit antibody binding but not to such a high degree as
CC   lactam bridge peptides. The lactam bridge containing peptides can be used
CC   in a combinatorial library of alpha-helical polypeptides each of 15-50
CC   amino acids, stabilised by at least one lactam bridge between non-
CC   adjacent residues and with a unique variation of amino acid residues in
CC   at least 3 positions. The library can be used to identify compounds that
CC   interact specifically with a particular macromolecular ligand, e.g.
CC   antibody (or fragment), receptor, ion channel, enzyme (or substrate)
XX
SQ   Sequence 11 AA;

  Query Match             35.3%;  Score 30;  DB 2;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          6 HFVKKHVH 13
              ||: :| |
Db          2 HFLVQHTH 9


RESULT 11
ADP83191
ID   ADP83191 standard; peptide; 12 AA.
XX
AC   ADP83191;
XX
DT   23-SEP-2004  (first entry)
XX
DE   ICAM-5 FGFR binding motif peptide SEQ ID NO:99.
XX
KW   protein interaction modulation; cell-surface receptor;
KW   receptor binding site; neuroprotective; vulnerary; antidiabetic;
KW   nephrotropic; gene therapy; fibroblast growth factor receptor; FGFR;
KW   receptor ligand; neural cell adhesion molecule; NCAM; solid tumour;
KW   neoangiogenesis; central nervous system disorder;
KW   peripheral nervous system disorder; postoperative nerve damage;
KW   traumatic nerve damage; impaired myelination; post-ischaemic damage;
KW   stroke; Parkinson's disease; Alzheimer's disease; Huntington's disease;
KW   dementia; multiinfarct dementia; sclerosis; nerve degeneration;
KW   diabetes mellitus; schizophrenia; mood disorder; manic depression;
KW   neuro-muscular connection; organ transplantation;
KW   genetic atrophic muscle disorder; traumatic atrophic muscle disorder;
KW   degenerative condition; acute myocardial infarction; angiogenesis;
KW   wound-healing; revascularisation.
XX
OS   Synthetic.
XX
PN   WO2004056865-A2.
XX
PD   08-JUL-2004.
XX
PF   18-DEC-2003; 2003WO-DK000901.
XX
PR   20-DEC-2002; 2002DK-00001982.
PR   03-MAR-2003; 2003DK-00000330.
XX
PA   (ENKA-) ENKAM PHARM AS.
XX
PI   Berezin V,  Albrechtsen M,  Bock E;
XX
DR   WPI; 2004-517671/49.
XX
PT   Modulating the interaction between at least two different proteins by
PT   providing a compound capable of interfering with the receptor and the
PT   polypeptide interaction and presenting the compound to the different
PT   proteins.
XX
PS   Claim 45; SEQ ID NO 99; 154pp; English.
XX
CC   The present invention describes a method for modulating the interaction
CC   between at least two different proteins, where one of the proteins is
CC   represented by a functional cell-surface receptor, or its fragment or
CC   variant or by a polypeptide having a binding site to the receptor, where
CC   at least a part of the binding site comprises at least one of the
CC   sequences given in SEQ ID NO:1 to 146 (ADP83093 to ADP83238), or
CC   fragments, or variants, or homologues of the sequences, or fragments or
CC   variants of the homologues, comprising (i) providing a compound capable
CC   of interacting with the receptor and/or polypeptide to interfere with the
CC   receptor and the polypeptide interaction, and (ii) presenting the
CC   compound of step (i) to at least two different proteins. The compound has
CC   neuroprotective, vulnerary, antidiabetic and nephrotropic activities, and
CC   can be used in gene therapy. The method is useful in modulating the
CC   interaction between at least two different proteins, where one of the
CC   proteins is represented by a functional cell-surface receptor, selected
CC   from the family of fibroblast growth factor receptors (FGFRs), and the
CC   other protein is the receptor ligand. A compound capable of modulating
CC   the interaction between at least two different proteins, where one of the
CC   proteins is represented by a functional cell-surface receptor, selected
CC   from the family of fibroblast growth factor receptors (FGFRs), and the
CC   other protein is the receptor ligand, can be used for treating normal,
CC   degenerated or damaged neural cell adhesion molecule (NCAM) presenting
CC   cells; solid tumour requiring neoangiogenesis; diseases and conditions of
CC   the central and peripheral nervous system, of the muscles or of various
CC   organs, e.g., postoperative nerve damage, traumatic nerve damage,
CC   impaired myelination of nerve fibres, post-ischaemic damage, e.g.
CC   resulting from a stroke, Parkinson's disease, Alzheimer's disease,
CC   Huntington's disease, dementias such as multiinfarct dementia, sclerosis,
CC   nerve degeneration associated with diabetes mellitus, disorders affecting
CC   the circadian clock or neuro-muscular transmission, and schizophrenia,
CC   mood disorders, such as manic depression; diseases or conditions of the
CC   muscles including conditions with impaired function of neuro-muscular
CC   connections, such as after organ transplantation, or such as genetic or
CC   traumatic atrophic muscle disorders; degenerative conditions of the
CC   gonads, of the pancreas such as diabetes mellitus type I and II, of the
CC   kidney such as nephrosis or of the heart, liver and bowel; for preventing
CC   cell death of heart muscle cells, such as after acute myocardial
CC   infarction, or after angiogenesis; for promoting wound-healing; for
CC   revascularisation; or for stimulating the ability to learn and/or the
CC   short and/or long-term memory.
XX
SQ   Sequence 12 AA;

  Query Match             35.3%;  Score 30;  DB 8;  Length 12;
  Best Local Similarity   45.5%;  Pred. No. 3.2e+02;
  Matches    5;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 GVYHFVKKHVH 13
              | || |  : |
Db          1 GTYHCVATNAH 11


RESULT 12
ADI32280
ID   ADI32280 standard; peptide; 13 AA.
XX
AC   ADI32280;
XX
DT   22-APR-2004  (first entry)
XX
DE   Cosmetic agent construct peptide linker #86.
XX
KW   cosmetic; keratin; dye; hair conditioner; combability improver;
KW   surfactant; amidoamine; betaine ester; esterquats; fatty alcohol;
KW   moisturiser; vitamin; provitamin; bounce improver; betaine; sugar;
KW   UV filter; polymer; silicone polyol; chitosan.
XX
OS   Synthetic.
XX
PN   WO2004000257-A2.
XX
PD   31-DEC-2003.
XX
PF   14-MAY-2003; 2003WO-EP005021.
```